Scuola Normale Superiore
Classe di Lettere e Filosofia

Tesi di Perfezionamento in Discipline Filologiche e Linguistiche Moderne

Individual Variation in Metaphor Comprehension and Processing

Relatore
Chiar.mo Prof. Pier Marco Bertinetto

Candidata
Simona Di Paola

Anno Accademico 2018-2019

Die Grenzen meiner Sprache

bedeuten die Grenzen meiner Welt.

The limits of my language mean

the limits of my world.

Ludwig Wittgenstein,

Tractatus Logico-Philosophicus, 1922

ABSTRACT

The primary objective of this thesis is to investigate the role of individual variation in metaphor interpretation. The scientific theoretical debate has traditionally treated metaphor interpretation from one and the same perspective which focuses exclusively on what the linguistic input is like (e.g., metaphoric vs. literal language) and which has disregarded up to now *who the processor is*. Experiments on metaphor understanding have been conducted on a similar vein, both in adulthood and development. Recently, however, a few preliminary studies have provided evidence in favour of the possibility that interpreting metaphor may change in terms of quality of the interpretation and/or underlying cognitive effort depending on some characteristics of the speaker. Yet, a systematic line of experimental research that informs us about the role of individual variation in metaphor interpretation seems essential to refine research on metaphor, but still it is missing. The main aim of this thesis is to shed light on this issue: understanding if and how metaphor processing varies depending on the speakers' characteristics. This was done by targeting two specific issues. First, by investigating which cognitive abilities scaffold the development of metaphor understanding in pre-school years. Second, by exploring the role of Education background and cognitive performance in adults' metaphor processing. Two experimental studies were thus designed and conducted which focused on individual variation in development (Study 1) and adulthood (Study 2).

Study 1 investigated if and how the development of metaphor comprehension is bound to the development of two cognitive abilities fundamental to metaphor comprehension, namely *Alternative Naming* (accepting two labels for the same referent) and *Analogy Perception* (detecting similarities across objects), whose role is still totally unexplored. These two abilities follow different developmental trajectories in pre-school years and may be crucial to explain the developmental path for metaphor. *Study 1* used a behavioural methodology and tested pre-schoolers in three separate tasks that assessed, respectively, metaphor comprehension, alternative naming and analogy perception. Employing a neuroscientific methodology, *Study 2* assessed the role of education background and individual cognitive performance in adults' metaphor processing. Participants with either an academic or a non-academic education read metaphorical and literal expressions while their

electroencephalogram was recorded. Additionally, a selected set of neuropsychological functions was tested.

Interesting results emerged from both studies. First, *Study 1*, Alternative Naming and Analogy Perception influence the development of metaphoric competence. By age 4, but not at 3, children's difficulties with alternative naming are solved and seem not to increase the cognitive demands imposed by metaphor. Conversely, analogy perception seems to hinder metaphor comprehension at both ages. Therefore, metaphorical development seems bound to alternative naming and analogy perception: depending on their developmental trajectories, these two abilities enhance, or impede, metaphor understanding in pre-school years. Second, *Study 2*, the electrophysiological activity associated to metaphor processing varies especially as a function of education background: the brain patterns for metaphor understanding changed substantially between adults with an academic education and adults with a non-academic education, thus suggesting that education background influences the cognitive activity underlying metaphor interpretation.

The results of both studies are discussed in detail in light of the available theoretical and experimental literature. Most important for the main purpose of this thesis, taken together, Studies 1 and 2 suggest that, both in development and adulthood, metaphor interpretation is associated to individual differences. Understanding metaphor does not seem the same for everyone, rather it changes in important respects depending on who the processor is. The variation depends on at least two individual sources: endogenous sources, such as differences in cognitive functioning, as well as environmental sources, such as education. This overall pattern raises non-trivial implications for theories of metaphor understanding and, more broadly, for mechanistic models of pragmatics that traditionally see pragmatic ability as an autonomous component of the mind.

Contents

Acknowledgements

Life reveals us only a few certainties. For as trivial as it might sound, one of these, for me, is that the most important achievements are hardly ever the easiest mountains to climb. Whoever works in academia would agree with me that doing a PhD is very much one of these cases. I could not have climbed this mountain without the precious support of some persons.

First, thank you to Prof. Pier Marco Bertinetto for choosing me as one of his PhD students and for nurturing my academic interests and supervising my work with expertise, patience and unfailing support. I would also like to thank Prof. Valentina Bambini. She has supervised my experimental work on adults and, in doing so, she provided me with the opportunity of becoming an experimental researcher. Thanks to this, I have experienced that in the fight between me and statistics, between me and programming, between me and some initially alien experimental machinery, I have been the winner rather than a bewildered "victim". If today statistics, numbers, programming and many other technical skills are my daily instruments with which I go on doing research, this is first and foremost thanks to the opportunities and challenges I was given during my PhD.

I am deeply grateful to Prof. Nausicaa Pouscoulous, who supervised my experimental research on development. Thank you for first introducing me to developmental pragmatics in a way that enchanted me when, a long time ago, I was a MA student at UCL. In fact, it's not by chance that one of my main research interests still is in the development of pragmatic abilities. Thank you for supervising my work during the PhD. I have incredibly benefited from the sharpness of your thought and the scientific rigour of your expertise and this allowed me to inherit a small, but precious, part of that expertise. Last but not least, thank you for encouraging and supporting me, both in research and in life. I suspect it will be difficult to pay you back for everything you have been doing, and still do, for me.

I owe a tremendous debt of gratitude to two special persons, Irene Ricci and Chiara Bertini. Both of you took precious care of me and I still remember this with admiration, nostalgia and affection. Thank you for the uncountable quantity of time you have spent helping me at making sense of statistics, functions, programming scripts, making the machinery work. Day after day, you dealt with my technical ignorance and, with enthusiasm and patience, you both shared with me your skilled expertise. You taught me to walk alone in the complex world of laboratory experiments. Also, it would be unfair if I would not thank you for being such great persons and friends. You protected me when in fact I needed; you supported and comforted me when I needed; you smiled with me and welcomed my ups and downs. So, every time, my ups got even greater, and my downs turned definitely more acceptable.

Impossible not to thank my personal friends and colleagues Giovanna Lenoci, Rosalba Nodari, Gianfranco Scuotri and Chiara Celata, without whom this doctorate would have never come to fruition. It was a pleasure to travel alongside all of you, to share the highs and lows of this marathon and take things with carelessness when carelessness was the best ingredient, and with seriousness when it was the case. Thanks to you, guys, I had a family when my suitcase brought me far away from my natural family. This is a rare privilege.

Finally, a special thank you goes to my family. My parents, Ettore and Sonia, and my brothers, Massi and Andrea, my grandparents and my uncles; your unconditional love and sense of shelter is a constant source of solace and resolve.

All these people have helped me staying afloat on an ocean that felt occasionally calm and clear and fun, occasionally rough and stormy. And they did so not even thinking twice. So, finally, I have learned to swim. Therefore, here is another certainty that I gained: no important achievement can be attained without someone who guides you in managing the right instruments for doing so. I am sincerely glad that so many people passed through my life with the only purpose of giving me the right instruments. I am lucky, definitely.

Introduction

The other side of the coin

There is almost no limit for human verbal communication to be successfully incisive, expressive, meaningful and creative at the same time. A happy mom's children can be *angels*, those of a stressed working woman can be *demons*. Your love affair can be a *prison,* a *nuclear war*, a *ride on the roller-coaster* or a *sweet hurricane*. Similarly, on a good morning, you can say to your husband "*I'm a coloured butterfly today".* Or, on a bad morning, "I'm *a tree on a windy and foggy winter".* In all these cases, what is being uttered is literally false and does not correspond to the actual state of affairs in the world: for instance, literally speaking, children are not supernatural creatures and a romantic liason is not a physical space nor it involves the use of nuclear weapons as prisons and nuclear wars respectively do. Yet, the examples above sound far from nonsensical. What they seem to communicate is actually a much richer message, one that conveys the speaker's meaning more vividly and deeply than what could have been expressed literally (cf., *I'm good today*) and one that somehow generates some effects on the hearer (e.g., imagery). In other words, when communicating, human beings creatively combine their linguistic and world knowledge and stretch the meaning of words well beyond literalness to powerfully convey their own intended meaning. Metaphor is one such way.

Metaphorical expressions are instances of non-literal uses of language in which something is talked about non-literally by means of something else, that is a different concept typically belonging to a different semantic category (e.g., human beings vs. butterflies). At the root of every metaphor is a *target* and a *vehicle* (e.g., *I* am a coloured *butterfly*), which are linked by the *metaphor ground*, that is an implicit relation between target and vehicle on which the metaphor is based. The starting point is a proposition whose literal meaning is false. The 'outcome' is a meaningful utterance in which the speaker's meaning becomes true and authentic.

The study of metaphor has been the object of an interest which dates back to Aristotle. Traditionally a matter for philosophers and literary critics, metaphor was generally treated as a mere rethorical device, a gimmick for poetry to decorate language, nice and graceful but not really necessary for communication (e.g., Aristotle, 350 BCEa, 350 BCEb; Richards, 1936; Cassirer, 1946; Black, 1962). It was only during the 1970s that metaphor placed at the crossroads of an increasing debate among linguists, psycholinguists and cognitive psychologists. On the one hand, with the birth of the pragmatics of language (Grice, 1975; Searle, 1979), it became unavoidable for linguists to address the issue of how a metaphor is interpreted since it is paradigmatic for the gap between what is said and what is meant by the speaker that characterizes ordinary language. On the other hand, it also became more and more clear that, far from being exclusive to the poet's pen, metaphor was pervasive in language use and was actually a primary means for ordinary conversation. Such an awareness gave birth to a long tradition of experiments on metaphor interpretation which was triggered by the need for any linguistic theory to account for our ability to produce and understand metaphor. Almost five decades later, the scientific debate on how metaphor is interpreted is still on air.

Within research in pragmatics, metaphor represents perhaps the most theorized figurative use of language. As mentioned above, what is said by the use of a metaphor - i.e., roughly, the semantic meaning of the words - is literally false; what is intended is meaningful and even true at some level. Consequently, the theoretical debate on metaphor comprehension has entirely focused on explaining how this gap is fulfilled. The main aim was to construct a theoretical model for the underlying interpretive process of a metaphor able to account for how the passage from (false) literal to (true) metaphorical meaning unfolds. To this purpose, theories on metaphor interpretation have developed their claims starting from the role they assign to literal meaning. Based on this, despite the inherent differences across theories, two main broad branches can be traced out. According to the original sequential models (e.g., Grice 1975/1989), metaphorical meanings are derived by understanding the literal meaning first, which is processed and rejected as false by the speaker and metaphor is inferred as part of the implicit meaning. Conversely, according to the models of parallel processing (e.g., Glucksberg, 2001; Wilson & Carston, 2007; Sperber &

Wilson, 2008), metaphorical meanings are automatically and directly accessed during on-line comprehension and literal meaning does not play any role but providing a hint for the interpretive process.

Accordingly, the leading issue for experiments on metaphor interpretation was to provide evidence for one or the other theoretical branch and look at how cognitively costly processing a metaphor is. The theoretical debate as sketched above has generated the general assumption that if metaphor interpretation passes through literal meaning, then extra-processing costs (e.g., slower reading times) should emerge in experiments comparing metaphoric to literal language comprehension. On the contrary, if no additional step is required (i.e., no rejection of literal meaning), then, no processing differences should occur during the interpretation of metaphorical and literal expressions. Results suggest a far from straightforward picture. In fact, on the one hand, some behavioural experimental studies find no processing differences between metaphorical and literal language and this suggests no additional cognitive costs for metaphor interpretation (vs. literal language) (see for example Glucksberg, Gildea & Bookin, 1982; but see also Noveck at al., 2001 for counterevidence and Glucksberg, 2003 for a review). On the other hand, studies using more fine-grained experimental methods such as electrophisiology and brain imaging costantly find increased brain activity for metaphor than literal language and thus provide evidence for extra costs required for metaphor comprehension (see Borhn et al., 2012 and Rataj et al., 2014 for reviews). In addition, it has been shown that the extra costs associated to metaphor are variable and are modeled by at least two main linguistic factors. First, the type of metaphor. Research has shown that interpreting conventional metaphors is less demanding in terms of processing costs than interpreting novel and/or literary metaphors (e.g., Arzouan et al., 2007b; Bambini et al., 2018). Second, the presence of a supporting context. In fact, there is experimental evidence suggesting that the costs associated to metaphor comprehension are reduced when metaphor is presented within context than when no context is provided (Bambini et al., 2016a).

Overall, this suggests that metaphor interpretation comes at a cost. Importantly, this cost is not always the same, but it changes depending on what the metaphor is and how it is presented. Finally, and consistently with this, metaphor comprehension is traditionally known to emerge late in development or adolescence (see Winner, 1988/1997 and Vosniadou, 1987 for reviews) or, at least, to emerge later than some other pragmatic skills such as for instance intention reading (e.g., Grosse, Schulze, Noveck, Tomasello & Katsos, 2011) and the derivation of other types of implicit meaning (e.g., scalar and relevance implicatures; Pouscoulous, Noveck, Politzer & Bastide, 2007; Schulze, Grassmann & Tomasello, 2013).

There is one main point that emerges from this short review of the literature on metaphor comprehension: the scientific debate has traditionally treated metaphor interpretation from one and the same perspective, one which focuses exclusively on what the linguistic input is and which has disregarded up to now *who the processor is*. In fact, both the theoretical positions shortly outlined above share one basic assumption: metaphor is interpreted according to one single processing mechanism which, passing or not through literal meaning, is always the same independent of who the processor is. In a similar vein, experiments on adults' metaphor understanding have tackled issues on metaphor processing related to the linguistic features involved in metaphorical language (e.g., metaphoricity vs. literalness; metaphor conventionality vs. novelty). Again, no attention has been drawn at who the processor is and how this could influence the costs for metaphor comprehension. As for development, the few available experimental studies have focused on when (i.e., at what age) *metaphor* understanding emerges, almost totally disregarding *how* it develops and what the developmental trajectory depends on. As a result, then, what is known about metaphor interpretation up to now only concerns half of the matter, that is what the object of the interpretation is like. Importantly, however, some experimental studies have provided evidence in favour of the idea that interpreting metaphor may change in terms of quality of the interpretation and/or underlying cognitive costs depending on some characteristics of the speakers. For example, people with a more efficient cognitive performance (e.g., better working memory ability) seems facilitated during the online comprehension of metaphor (e.g., Olkoniemi et al., 2016; Chiappe & Chiappe, 2007). In addition, results of a few experimental studies favour the possibility that education-related factors such as the number of years of schooling and the type of programme attended (i.e., humanities vs. non-humanities) may enhance the interpretation of figurative language in young adults (Burkett & Goldman, 2016; Peskin, 2010) and that older adults with a higher education appreciate metaphor better than older adults with a lower education (Champagne-Lavau et al., 2012). Finally, one experimental study showed that executive functions play a prominent role in metaphor comprehension during adolescence (Carriedo et al., 2016). Taken together, these data point to the possibility that metaphor interpretation/processing might actually be even more complex than currently known since it may vary not only depending on what the metaphor is like (e.g., novel vs. conventional or familiar vs. unfamiliar), but also depending on who the processor is. If this was the case, then, what is known up to now about metaphor interpretation would only represent one side of the coin. In fact, except for the experimental studies briefly mentioned above, a systematic line of experimental research that investigates the role of individual variation in metaphor interpretation seems essential to refine and advance research on metaphor comprehension but still it is missing.

The main aim of this dissertation is to investigate experimentally the role of individual variation in metaphor interpretation and shed light on the other (unexplored) side of the coin, that is understanding if and how metaphor processing varies depending on the speakers' characteristics. In the present dissertation, this was done by targeting two specific issues. First, by investigating which cognitive abilities, if any, scaffold the development of metaphor comprehension in pre-school years. Second, by exploring the role of Education type and cognitive performance in adults' metaphor processing. Two experimental studies were thus conducted to cast light on these issues and focused on the role of individual variation in development and adulthood, respectively.

Study 1 assesses if and how the development of metaphor comprehension is bound to the development of two cognitive abilities fundamental to metaphor comprehension, that is Alternative Naming (i.e., accepting two labels for the same referent) and Analogy Perception (i.e., detecting similarities across entities), whose role is still totally unexplored. Alternative Naming is involved in metaphor comprehension because, in a metaphor, the same entity is referred to using two linguistic labels, the literal label and a figurative, more unconventional one (e.g., *humans* are *butterflies*). Analogy Perception is involved because in order to figure out the metaphor ground it is first and foremost necessary to identify some shared properties between the metaphoric components (e.g., *what properties are shared between humans and butterflies*?). These two abilities follow different developmental trajectories in pre-school years and may be crucial to explain the development of metaphoric competence as difficulties in one or both of them may impose additional demands linked to children's general cognitive development, hence slowing down the development of metaphor understanding. Thus, Study 1 aims at teasing apart the contribution of Alternative Naming and Analogy Perception in the development of metaphor comprehension. To this purpose, three- and four-year-olds were tested within the same picture-matching scenario in three separate tasks which assessed separately metaphor comprehension, alternative naming and analogy perception. The three tasks were all inspired by previous experimental studies and adapted such that they could be administered with the same general procedure. This resulted in a newly developed experimental paradigm.

Study 2 assesses the role of Education background and individual cognitive performance in adults' metaphor processing. The main aim of Study 2 was to understand if and to what extent variation in adults' metaphor processing may depend on - and be at least partially explained by - a socio-demographical factor such as Education type and/or the speakers' cognitive efficiency. In order to do so, two groups of participants were recruited, with either an academic or a non-academic education respectively. The two groups read metaphorical and literal sentences while their electroencephalogram was recorded. Additionally, a battery of neuropsychological measures was administered and this included standardized tests for Verbal Working Memory, Verbal Fluency, IQ, Autism Quotient and Exposure to reading.

Study 1 is reported in Chapter 1. Study 2 is reported in Chapter 2. Finally, some overall conclusions are drawn at the end of the present dissertation.

Chapter I

Study 1 - Metaphorical Developing Minds: The role of multiple factors in the development of metaphor comprehension

1. Introduction

You are my lovely marshmallow. Who hasn't ever heard a mum speaking like this to her beloved child? While the mum's communicative intention in the use of such a novel metaphor is pretty clear (i.e., communicating love and affection to the child), the issue of whether - and when - a young child is able to grasp metaphorical meanings, and if not why, is much more debated by researchers on the development of metaphoric comprehension.

Metaphors are instances of language use in which something is described by means of something else, they are composed of a target (e.g., 'you' referring to the baby) and a vehicle (e.g., the marshmallow) used to communicate something about the target and usually belonging to a different semantic category. Typically, in metaphorical expressions, what is uttered is not literally true: babies are not candies, they are not rubbery and are not made of sugar, to start with. In order to grasp the meaning of a metaphor, speakers have to jump from false literal to true metaphorical meaning and this makes metaphors paradigmatic uses of language where what is meant does not correspond to what is literally said.

The path to bridge this gap would be a sterile land without an appropriately developed metaphorical competence – i.e., a set of linguistic and cognitive abilities to comprehend and produce metaphors. When and how metaphorical competence develops?

Metaphor understanding is traditionally thought to emerge fairly late in childhood and even adolescence. In addition, its developmental trajectory tends to be delayed with respect to the development of general pragmatic and inferential skills. Why is this so? The main aim of Study 1 is to investigate on some of the sources of such difficulty.

1.1 The development of metaphor comprehension

The developmental time-course of metaphor comprehension is a largely debated issue. Classical experimental studies from the '60s, '70s and '80s suggest that young children have difficulties with the genuine appreciation of metaphors and that this ability develops fairly late in adolescence. Recognizing the non-literal dimension of language would start at 8 years of age, but children fully master metaphors only later while interpreting them literally before age 8 or 9 (e.g., Billow, 1975; Cometa & Eson, 1978; Asch & Nerlove, 1960; Winner et al., 1976; see Winner 1988/1997 and Vosniadou, 1987 for reviews).

This view has been argued against recently on the basis that rather than indicating poor pragmatic skills, these findings could be due to confounding variables in the experimental designs, such as the use of metalinguistic tasks, age-inappropriate lexical and conceptual knowledge and not sufficiently controlled materials (Pouscoulous 2014, 2011). Indeed, the use of more child-friendly experimental settings controlling for these factors, lowers at 5 (Wagonner & Palermo, 1989) or even 3 (Pearson, 1990; Pouscoulous & Tomasello, 2011) the age at which signs of metaphorical competence are visible. These studies suggest that the pragmatic skills necessary for metaphor understanding are likely to be already present in pre-school years, when elicited under more fine-tuned experimental manipulations. Such a trend is fully consistent with recent findings showing that the ability to derive other pragmatic inferences, such as scalar and relevance implicatures or presuppositions, emerges early in development (Grosse, Schulze, Noveck, Tomasello & Katsos, 2011; Kastos & Bishop, 2011; Pouscoulous, Noveck, Politzer & Bastide, 2007; Schulze, Grassmann & Tomasello, 2013; Berger & Höhle, 2012).

Still, other recent studies bring to light two interesting patterns. First, metaphor understanding improves as a function of age (Declercq et al., 2010). Second, there seems to be a stark improvement from 3-to-4, with 3-year-olds either performing at chance (Özçalişkan, 2005) in a metaphor comprehension task and/or significantly worse than 4- and

5-year-olds (Deamer & Pouscoulous, Experiment 1, submitted). The above studies used a variety of experimental tasks, including multiple choice paraphrases, story comprehension and picture-selection tasks. The difference in task demands might be responsible for younger children's different performance across studies (e.g., below or above chance). In any case, the overall emerging developmental trend is quite homogeneous: children show difficulties with metaphor understanding during pre-school years and tend to interpret metaphor literally. Moreover, although research on the development of different figurative uses has not been extensively carried out yet, there is evidence in favour of the possibility that young children are less proficient and exhibit greater difficulties with metaphor than other non-literal uses of language such as hyperbole (Deamer & Pouscoulous, Experiment 2, submitted) and metonymy (Rundblad & Annaz, 2010; Falkum et al., 2017).

Overall, then, if on the one hand age-inappropriate experimental paradigms undoubtedly impede children's grasping of a metaphor and are a source of difficulty independent of children's metaphorical abilities *per se*, on the other hand they are likely not to be the only source of difficulty. As suggested by recent studies using more child-friendly paradigms, pre-schoolers' ability to overcome literal language when tapping into metaphor seems more developed than traditionally thought, but still it is restricted. Where children's metaphoric difficulties stem from, beyond external factors (e.g., experimental designs), is one of the main missing pieces of the puzzle in developmental research. As featured in the next section, metaphor comprehension involves multiple higher-order cognitive skills. To this purpose, looking at the development of the abilities that scaffold metaphoric competence may help identifying some of children's sources of difficulty.

1.2 Metaphor processing: multiple factors and sources of difficulty

Pragmatic theories have offered several accounts on metaphor comprehension. Although nowadays the original idea of a sequential interpretive process based on understanding the literal meaning first (e.g., Grice 1975/1989; Searle, 1979) is largely rejected and models of parallel processing are favoured (e.g., Glucksberg, 2001; Wilson & Carston, 2007), there is much less consensus on the nature of the process involved in the passage from literalness to metaphoricity: some explain metaphor interpretation in terms of a mapping of properties (of the target and vehicle) belonging to two different domains (Lakoff & Johnson, 1980/2003); some others see metaphors as implicit comparison statements and explain its interpretation in terms of a comparison of the properties of the vehicle and the target until a final matching is achieved (e.g., Gentner et al., 2001; Wolff & Gentner, 2000); some others have claimed that metaphors are categorization statements (e.g., Glucksberg, 2001) and provide a deflationary account according to which, together with other figurative expressions, metaphor is an instance of loose use of language processed through the online construction of occasion-specific concepts via broadening and/or narrowing of the lexically encoded concepts (Carston, 2002; Wilson & Carston, 2007).

Whatever the interpretive mechanisms are, all accounts of metaphor interpretation presuppose two basic features. First, the target is described by use of the vehicle, that is generally an entity belonging to a different category. This unavoidably involves dual reference: in a metaphor, the target is referred to both as the target and the vehicle. Thus, the same object is talked about with two different names in the same communicative context (e.g., a baby is a *baby* and a *marshmallow* at the same time). Second, category violation calls for some kind of mental operations on the properties of the vehicle to be meaningfully applied to the target. Whatever the nature of such operations is - either mapping or broadening or implicit comparison - they all presuppose first and foremost the identification of relevant properties needed to recognise the metaphor ground, that is the implicit relation between target and vehicle.

When turning to processing, the above features involve two cognitive abilities not exclusive to but fundamental for metaphoric comprehension, namely Alternative Naming, i.e., accepting two labels for the same referent, and Analogy Perception, i.e., detecting similarities across objects (see Pouscoulous, 2014; Rubio-Fernández & Grassmann, 2016). Dual reference involves Alternative Naming since it presupposes the ability to accept that one and the same entity can take two linguistic labels within the same communicative context, a literal and a more unconventional figurative one (i.e., given that babies are not marshmallows can a baby *be called* 'baby' and 'marshmallow' at the same time?). Likewise, the identification of relevant properties to figure out the metaphor ground presupposes the ability to detect similarities across objects and entities (i.e., what properties are shared by babies and marshmallows?), which leads to the involvement of analogical skills[1].

To the best of my knowledge, the only study addressing this issue is Rubio-Fernández & Grassmann (2016). Based on the above assumptions, they tested 3-and-4-year-olds' ability to integrate implicit spatial reasoning and alternative naming within a referential task which involved literal language only. It was found that both age groups exhibited remarkable difficulties in the condition where both skills were required, with 3-year-olds performing at chance and 4-year-olds performing significantly worse than in the baseline condition, where implicit spatial reasoning only was elicited. The authors interpret their findings as suggesting that pre-schoolers' metaphor comprehension might be compromised by their difficulty with alternative naming.

These results constitute first evidence in support of the idea that young children's metaphoric difficulties might be bound to their general cognitive development and that the integration of alternative naming and analogy perception plays a role in the development of metaphor comprehension. However, in Rubio-Fernandéz & Grassmann (2016), children were tested in a task which involved literal language only, hence it is still unclear whether the two abilities at stake play a prominent role for metaphor comprehension since more direct comparisons are still missing. In addition, as the authors themselves acknowledge (p. 940), analogy perception was assessed only with respect to implicit spatial relations, that is a very specific and basic subtype of analogical transfer which children master from very early on (Gentner, 1977). Further research is needed to better identify the role analogy perception plays in metaphor development, especially with respect to children's mastering of physical relational features which include but are not limited to spatial relations. Last but not least, Rubio-Fernandéz & Grassmann (2016) provides direct evidence about pre-schoolers' ability of *integrating* alternative naming and analogy perception. Again, further research is needed to tease apart the contribution of each of the two abilities in the development of metaphor understanding.

[1] Note that the involvement of analogy perception is independent of an understanding of metaphor as a theory of analogy or not (see Gentner, 1988; Bowdle & Gentner, 2005 for an account of metaphor understanding massively relying on analogy). In fact, whatever the theory – and the role analogy plays – it would be difficult to imagine how one could understand or produce metaphor without any involvement of analogical skills.

To summarize, during their developing language experience, when tapping into metaphor, children unavoidably tap into alternative naming and analogy perception too. These two cognitive abilities follow different developmental trajectories and, already for this reason, are likely to impede pre-schoolers' grasping of a metaphor and delay the development of metaphoric competence: difficulties in one or both of them may impose additional demands for metaphor processing linked to pre-schoolers' general cognitive development.

1.2.1 Alternative Naming

Alternative naming is the metalinguistic ability to assign a second label to a given entity and to accept that one and the same entity can take more than one label. Children as young as 2 years of age do learn a second label for an object (e.g., they know a 'mouse' is an 'animal' too, see for example Grassman & Tomasello, 2009) but it seems that alternative naming is problematic for pre-schoolers when it occurs within the same communicative circumstance, that is when the same object is given two labels in the same context (e.g., Doherty & Perner, 1998).

Some experimental studies have shown that accepting a known, conventional word for a known object is somewhat more demanding for children until at least their 4th year of age and that this is even more so when a referential pact is violated. Matthews, Lieven & Tomasello (2010) tested 3-and-4-year-olds in a referential communication task in which the first conventional label (e.g., "car" for a toy car) was used by one experimenter (same term/original partner condition) and then an alternative label (e.g., "truck" for the same toy car) was used either by the original experimenter (new term/original partner condition) or by a new one (new term/new partner condition). Children of both age groups took the correct object (e.g., the toy car) most of the times, but three important results emerged in this study. First, children were significantly slower in the new term condition compared to the same term one. Second, the difference reaction times were significantly greater for the original partner condition (i.e., same context) than for the new partner one - where no referential pact was previously established hence violated. Finally, such difference score was significantly affected by age with greater difference reaction times for 3- than 4-year-olds. These results are also consistent with classical studies on alternative naming which used (known and fully consistent) synonyms (e.g., rabbit - bunny) or superordinate categories (e.g., rabbit - animal) as second labels presented within the same context and found that children's ability to accept a second label in the same context positively increased as a function of age, with severe problems for 3-year-olds and a stark improvement from age 4 (Doherty & Perner, 1998; Perner et al., 2002).

Overall, the existing experimental literature brings to light two main points. First, children have difficulties accepting that two labels are used for the same entity in the same communicative context. Second, they tend to overcome such difficulties from age 4. Perner and colleagues claim that, in tasks where alternative naming occurs within the same communicative context, young children's problems are ascribable to their insufficiently developed ability to take an alternative perspective, that is to understand that the same thing - (individual, event, object, etc.) can be described and mentally represented differently from different perspectives (Perner et al., 2002; Perner et al., 2003; Perner et al., 2007; Perner & Leahy, 2016). This latter ability is found to strongly correlate with theory of mind (e.g., Doherty & Perner, 1998), that is known indeed not to develop before age 4, when children pass the standard false-belief task (Baron-Cohen, 1995).

Alternative naming might impose additional cognitive demands on young children's metaphoric interpretation: metaphor typically involves the use of an unconventional (i.e., not literal) label to refer to a known object (e.g., marshmallows for babies), which also implies taking two different perspectives on the same entity within the same communicative context. Although alternative naming and metaphor differ in many respects, it is definitely hard to understand a metaphor without accepting that two known labels - a literal and a figurative/unconventional one - are being used for the same referent in the same context. Alternative naming might represent a source of difficulty for children's metaphorical competence and, considering its developmental trajectory, this should be so at age 3 but presumably not later than age 4.

1.2.2 Analogical Reasoning

Analogical reasoning is the ability to transfer one mechanism from one known entity/system to another (less known) entity/system. It is first and foremost enhanced by the ability to identify the relational structure shared by two or more analogical terms, that is the fact that two or more objects are linked by the same relation despite being very different on surface level. Such constraint is known as the relational similarity constraint (Goswami, 1991) and is considered as the hallmark of analogical reasoning (see Goswami, 2001 for a review).

Analogical reasoning has been extensively studied in development and is known to be present from at least 3-and-4 years of age, although some rudimentary forms of analogical reasoning capacities have been found to develop much earlier in infancy from 6 months of age onwards (e.g., Wagner et al., 1981; Chen et al., 1997).

The standard test used to assess its development is the item analogy task, administered in a pictorial version as a game about matching pictures to test young children (Goswami, 1989; Goswami & Brown, 1989, 1990). This paradigm requires the use of the relational similarity constraint for successful completion of an analogical sequence: 2 items A and B are presented to the child, a 3rd item C is then presented and the child is required to generate a D term (or select the correct D picture) that has the same relation to C as B has to A (e.g., "Cat is to kitten as dog is to?" with "puppy" being the correct solution term). Children have been shown to succeed this task from age 3, provided the child is familiar with the relations used in the materials. For example, Goswami & Brown (1989) tested 3- 4- and 6-year olds in a picture-sequencing task and asked children to choose the picture of the correct D terms from a range of 5 alternatives. They found that children of all age groups performed above chance, with 6-year-olds reaching ceiling levels. Interestingly, 3- and 4-year-olds' most frequent errors were either a picture of the correct object undergoing the wrong physical change or a picture of the wrong object undergoing the correct change, which the authors interpret as (unsuccessful) attempts at reasoning by analogy.

Overall, then, analogical reasoning is relatively early-developing. However, there is evidence suggesting that pre-schoolers' analogical abilities depend on the level of complexity involved in the analogy task (e.g., Richland et al., 2006; Tunteler & Resing, 2002, 2007). In addition, relevant indication with this respect is also provided by Goswami & Brown (1989), who found that when the analogical problem is explained to the child, (i.e., the salient relation and the analogical transfer are made explicit), children as young as 3 succeed a task where, in order to select the correct picture, it is necessary to identify the relevant property and apply it to the correct object. Or, at least, they appear to be able to reason about higher-order relations: they recognize the salient relation for an analogical sequence to be successfully completed, even though they do not always apply it correctly.

Conversely, when the analogy problem is inducted (i.e., presented in a more implicit form), even 4-year-olds exhibit greater difficulties.

These differences may be relevant indeed with respect to pre-schoolers' metaphorical abilities. In fact, in order to understand a metaphor, it is necessary to grasp the metaphor ground where, importantly, the relation between target and vehicle is only implicitly conveyed (see also Rubio-Fernández & Grassmann, 2016). Thus, even though pre-schoolers possess at least basic analogical abilities, still their perception of analogical relations might not be sufficiently developed to figure out the relevant implicit property and apply it to the correct object. In this respect, analogy perception - at the level of complexity as involved in metaphor - might still represent a source of difficulty in preschool years.

1.3 The present study

The main goal of this study was to extend Rubio-Fernandéz & Grassmann's (2016) findings and tease apart the contribution of Alternative Naming and Analogy Perception in the development of metaphor understanding during pre-school years. Assuming that the understanding of metaphorical language involves multiple higher-order cognitive abilities, it was hypothesized that its developmental trajectory might in turn be influenced - and at least partially explained - by the development of alternative naming and analogy perception. I reasoned that some of pre-schoolers' difficulties in metaphor interpretation might be due to difficulties with alternative naming and analogy perception, which makes these two abilities possible predictors for the development of metaphor understanding.

To the best of my knowledge, this study is the first directly investigating the role of alternative naming and analogy perception in pre-schoolers' understanding of a metaphor. This was done by developing a new experimental procedure. Three- and 4-year-olds were tested in a metaphor comprehension task, an alternative naming task and an analogy perception task. The three tasks assessed separately pre-schoolers' metaphorical, alternative naming and analogical abilities but, importantly, they were built within the same experimental scenario based on a picture-selection general procedure. The three tasks were all inspired from previous experimental works: the metaphor task was adapted from Deamer & Pouscoulous (submitted); the alternative naming task was based on Matthews, Lieven & Tomasello (2010); the analogy perception task was inspired from Goswami &

Brown (1989). These were all adapted such that the three abilities could be tested using the same type of paradigm. Children were asked to move some pictures around a grid following the experimenter's instructions. In the metaphor task, which included a metaphoric and a literal condition, the experimenter referred to the pictures either metaphorically or literally. In the alternative naming task, the target picture was referred to either using the same previously uttered label or a new label. In the analogy perception task, children were asked to complete a sequence of two pictures by choosing one more picture from a range of alternatives (see procedure section for further details on each task).

Overall, using the same type of paradigm is crucial because this considerably contributes to reduce the discrepancies between experimental methods tackling the various abilities at stake. Testing children's alternative naming and analogy perception abilities within the same experimental scenario is thus two-fold: on the one hand, it allows us to investigate on each of these abilities *per se*, on the other hand it allows us to detect any correlation between their developmental trajectories and statistically verify whether alternative naming and analogy perception might significantly predict the development of metaphoric competence using large-scale regression models. Last but not least, assessing children's metaphorical abilities within a task which includes a baseline literal condition is crucial too because it allows pre-schoolers' difficulties with metaphor understanding - if any - to emerge more clearly.

For metaphor comprehension, it was predicted that pre-schoolers would have greater difficulties with metaphoric language than with literal expressions. With respect to alternative naming and analogy perception, I predicted that pre-schoolers' proficiency with metaphor understanding should increase as a function of their proficiency in alternative naming and analogy perception. Both abilities at stake should act as enhancing or impeding factors depending on their developmental trajectories. Based on previous developmental literature, alternative naming should represent a source of difficulty for metaphor comprehension at age 3, but it should be resolved by age 4. Based on previous literature on analogy, I predicted that by age 4, although not necessarily at 3, children should at least show evidence of correctly identifying the salient property/relation needed to solve an analogy presented by induction (i.e., in a more implicit form), even though this might not be sufficient for the understanding of a metaphor. Thus, when involved at a higher level of complexity, analogy perception might represent a source of difficulty too.

2. Methods

2.1 Participants

Thirty-nine monolingual Italian speaking children participated in the experiment: 19 3-year-olds (M=3;4, range 3;1 to 3;11, 11 girls) and 20 4-year-olds (M=4;5, range 4;1 to 4;11, 9 girls). They were recruited from and tested in one nursery school and pre-school in a middle-sized city in Italy. Informed consent was obtained from parents or guardians of all participating children prior to the beginning of the experiment. This study was approved by the Research Ethics Committee of University College London (Project ID number: 3150/004) and the Scuola Normale Superiore.

As a general criterion, children who did not perform correctly in the familiarization phase of the grid game after being corrected more than twice were excluded from the analyses. Based on this criterion, one 3-year-old child was excluded from the analyses. No other children were excluded from the analyses.

2.2 Materials

For the metaphor task, the alternative naming task and the analogy perception task a 100x79 cm wooden grid was used. The grid was composed of 16 slots equipped with a magnet. The pictures used in each task were 22x15 cm magnetic plastified pictures children could attach on and detach from the grid.

2.2.1 The Metaphor comprehension task

Eight pairs of novel metaphors and corresponding literal expressions were used to refer to the target objects shown in the pictures during the metaphor session of the grid game (see Table 1). All metaphoric and literal utterances were expressions of the form [The X with the Y]. The vehicle of the target expression always referred to a salient property of the target picture: in the metaphor condition, the salient property was referred to metaphorically (e.g., *the glass with the antennae* for a glass with two straws), while in the literal condition it was referred to literally (e.g., *the glass with the straws*). Only metaphors based on physical similarities between target and vehicle were used to ensure the material was age-appropriate with respect to children's world knowledge and conceptual development (see Winner 1988/1977 for children's perception of abstract/psychological relations). Three of the 8 metaphors used were taken from Pouscoulous & Tomasello (2011) (i.e., *The car with the backpack*; *The carrot with the hair; The tower with the hat*), while the others were specifically created for the present study using words Italian children as young as 2 years of age are known to produce (Caselli & Casadio, 1995).

The material was administered with a Latin-square design: the 8 pairs were randomly divided into two lists such that all children saw both conditions, but the same child never heard a metaphor and its literal counterpart. Children were randomly assigned to one of the two lists, each containing 4 metaphorical and 4 literal scenarios.

Table 1. List of the metaphorical and literal expressions used in the Metaphor Task. A literal translation in English and the original expressions in Italian are provided.

	Metaphor	**Literal**
1.	The car with the backpack *Ita: La macchina con lo zaino*	The car with the box on the top *Ita: La macchina con la scatola nel tetto*
2.	The kitten with the socks *Ita: Il gattino coi calzini*	The kitten with the white paws *Ita: Il gattino con le zampe bianche*
3.	The carrots with the hair *Ita: Le carote coi capelli*	The carrots with the leaves *Ita: Le carote con le foglie*
4.	The bottle with the big belly *Ita: La bottiglia col pancione*	The round bottle *Ita: La bottiglia rotonda*
5.	The tower with the hat *Ita: La torre con il capello*	The tower with the pointy roof *Ita: La torre col tetto a punta*
6.	The tree with the arms *Ita: L'albero con le braccia*	The tree with the branches *Ita: L'albero con i rami*

7.	The glass with the antennae *Ita: Il bicchiere con le antenne*	The glass with the straws *Ita: Il bicchiere con le cannucce*
8.	The mobile phone with the coat *Ita: Il cellulare col cappotto*	The mobile phone with the cover *Ita: Il cellulare con la custodia*

In each trial, children saw 3 pictures, which were the same in both conditions: the target picture, which illustrated the target object the experimenter referred to either metaphorically or literally, and 2 controls (see Figure 1 for an example and Appendix A for all the experimental material). The target picture always displayed the target object with the salient property described either metaphorically or literally (e.g., a glass with two straws for *The glass with the antennae/glass with the straws* trial). Control picture *Target-Literal* always displayed the target object deprived of the salient property (e.g., a glass with no straws). Control picture *Vehicle-Literal* was a literal competitor always displaying both the entities of the metaphor target and vehicle: one of the two entities represented the metaphor target without the salient property, while the second entity was a person or object literally possessing the attribute referred to by the metaphor vehicle (e.g., the picture of a glass and a girl wearing an antennae-headband for *The glass with the antennae*). The two controls were selected to better classify children's interpretation of the metaphor: I reasoned that (i) if children pick up the *Target-Literal* distractor, their selection might be guided by object preference or attention only to *The X* component of the metaphor; (ii) if children pick up the *Vehicle-Literal* distractor, then this would be a clear indication of a genuinely literal interpretation of the metaphor; and consequently (iii) if children pick up the target picture, this would indicate genuine metaphor understanding.

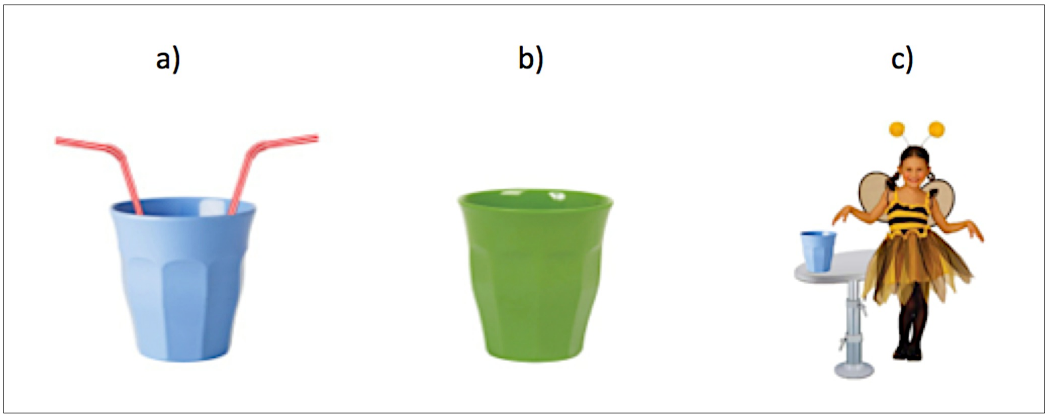


Figure 1. Picture alternatives for the metaphor item *The glass with the antennae* and the literal item *The glass with the straws*: (a) target picture, (b) *Target-Literal* distractor, (c) *Vehicle-Literal* distractor.

2.2.2 Naming -and-pointing picture book

A naming-and-pointing picture book was created to assess children's knowledge of the vocabulary used in the metaphor task. The book was adapted from Pouscoulous & Tomasello (2011) and consisted of a comprehension and production part. Four pictures were used to test the vocabulary of each metaphorical trial (Figure 2): picture 1 showed the literal referent of the metaphor vehicle (e.g., the antennae of a snail for *The glass with the antennae*) and was inserted to assess whether children could comprehend/produce a literal label for the target referent; picture 2 displayed the intended referent of the metaphor vehicle (e.g., a straw into a glass of juice) and was inserted to assess any eventual case of overextension of a term; pictures 3 and 4 were distractor items. Different pictures were used for the comprehension and production tasks.

Figure 2. Sample items used in the comprehension (left) and production (right) part of the naming-and-pointing picture book for the metaphor *The glass with the antennae*, for which the words 'antennae' and 'straw' were tested*.*

2.2.3 The Alternative Naming Task

Thirteen pictures were used in the alternative naming session, 8 target pictures and 5 filler items. Among the set of target pictures, 4 were used for the Same Term Condition (S.T.) and 4 for the New Term Condition (N.T.). In the S.T. condition, the literal and conventional label was always referred to the target picture (e.g., *Backpack, Ita. Zaino,* for the picture of a backpack). In the N.T. condition, 4 pairs of literal expressions were used to refer to the target pictures: the first term was the literal and conventional label for the target referent (e.g., *Lollipop, Ita. Leccalecca,* for a picture displaying a lollipop), while the second term was an alternative label belonging to the same semantic category of the first label and which could be felicitously used in Italian to refer to the target picture (e.g., *Candy*, *Ita. Caramella,* for a picture displaying a lollipop). The expressions used to refer to the 8 critical pictures both in S.T. and N.T. are presented in Table 2. Two of the N.T. pairs were taken from Matthews et al. (2010) (i.e., Truck/Car; Book/Story).

Table 2. List of the target expressions used in the Alternative Naming session of the Grid Game for the Same Term (S.T.) and the New Term (N.T.) condition *(original expressions in Italian).*

S.T.	**N.T.**
Kitten (*Micio*)	Truck/Car (*Camion/Macchina*)
Boots (*Stivali*)	Book/Story (*Libro/Favola*)
Spaghetti	Lollipop/Candy *(Leccalecca/ Caramella)*
Backpack (*Zaino*)	Colours/ Pencils *(Colori/Matite)*

2.2.4 The Analogy task

There were 4 trials in the analogy task, each consisting of a sequence of 2 pictures (A & B) to be completed with a C picture, chosen from a range of 3 alternatives.

The analogies were created according to the following criteria. First, surface similarity (i.e., the similarity shared by two or more objects such as for example shape and colour) is a factor known to bias children's analogical reasoning (e.g., Gentner, 1989). Thus, all analogies were constructed respecting the relational similarity constraint (Goswami, 1991) in the absence of any surface similarity. That is, the objects constituting the analogical sequence were not similar in shape, colour or any other physical property, but they were linked to each other by the same relations. Second, the analogies were based on simple relations young children are highly familiar with (e.g., simple causal relations such as opening, see Goswami, 2001, 1990). Third, the analogies were based on an appropriate world knowledge for children as young as 3 years of age (i.e., selection of entities/situations children commonly experience in real-life). Each analogical trial consisted overall of a 3-term-comparison (A, B & C) and involved one of the following relations: (i) *Housing relations* (i.e., A: nest; B: fish bowl; C: humans' house); (ii) *Family relations* (i.e., A: penguin family; B: bear family; C: human family); (iii) *Open relations* (i.e., A: open gift box; B; open bottle; C: open cookies box); and (iv) *Animal flying relations* (i.e., A: flying butterfly; B: flying bird; C: flying bee). For each trial, children were presented with terms A and B (e.g., a flying butterfly and a flying bird for the *Animals that fly* trial) and were asked to complete the sequence by picking up a picture out of 3 which best completed the pattern (i.e., the C term).

The 3 picture choices were 1 target item and 2 control items, whose construction was adapted from Goswami & Brown (1989). The target item was a picture displaying the relevant property applied to the relevant object. For example, in *The animals that fly* trial, the A and B terms were a flying butterfly and a flying bird and the target C item was the picture of a flying bee. Control item 1 - *Correct property/wrong object* - was a picture showing the relevant property but applied to the wrong object (e.g., a helicopter), while Control item 2 - *Wrong property/correct object* - was a picture depicting an irrelevant property (for the purpose of the analogical sequence) applied to the right object (e.g., a sleeping bee). See Figure 3 for an illustration of the pictures used in one trial and Appendix A for all the material used in the Analogy Perception task.

Note that in the original paradigm, Goswami & Brown (1989) used 4 control items to assess whether children solved the analogies by reasoning about higher-order relations (what they label *items E & F*, pag. 76) or by associative mechanisms based on appearance similarity or thematic/category match (*items G & H*, pag. 76). In the error analysis, they found that children's most frequent errors both at ages 3 and 4 were either a picture showing the wrong object undergoing the correct physical change or a picture showing the correct object undergoing the wrong physical change. This was interpreted by the authors as evidence that, even when they are not able to solve an analogical problem, pre-schoolers can still reason about higher-order relations.

The analogy perception task in the present study aimed at assessing the extent to which analogical perception is developed in pre-school years because this might be crucial for the development of metaphor understanding. This is why the alternative choices were restricted to 1 target item and the 2 control items *Correct property/wrong object* and *Wrong property/correct object* adapted from Goswami & Brown (1989, items E & F). In order to figure out the implicit relation between the target and vehicle of a metaphor (i.e., metaphor ground), it is necessary to both identify a relevant property of the vehicle and to apply it to the target of the metaphor. In other words, identifying a relevant property might not be enough on its own to fully recover the metaphor ground. To this purpose, children's selection of one or the other control can be very informative. In both cases, an ability to reason about higher-order relations is presupposed but, crucially, while the selection of the *Wrong property/correct object* item (e.g., the sleeping bee) would indicate a substantial lack of analogical abilities because of the absence of the salient relational property, the selection of the *Correct property/wrong object* (e.g., the helicopter) would indicate that children are able to identify the salient property, but their analogical skills are not yet developed enough to fully complete the elicited analogy.

Fifteen native speakers of Italian, aged between 20 and 40, were tested with the material used in the metaphor, alternative naming and analogy tasks. They chose the target picture 100% of the time, indicating that the experimental material was understandable and suitable. In addition, 5 9-year-olds were tested with the material used in the analogy task. They chose the target item 100% of the time.

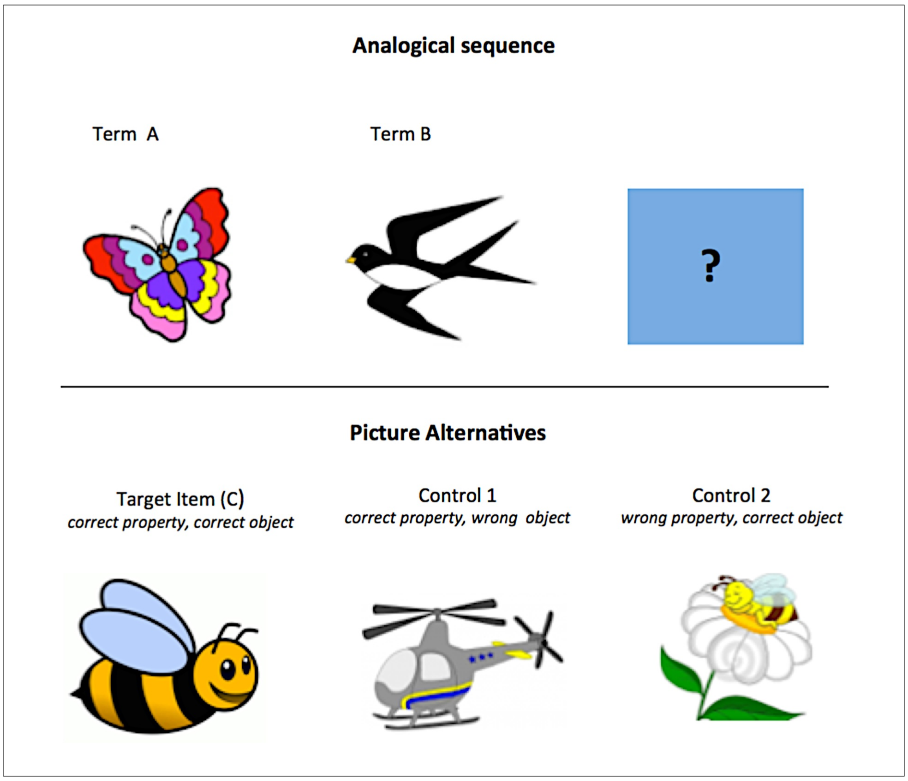


Figure 3. Example of the material used in the Analogy Perception task. Here, the material for the *Animals that fly* trial is illustrated.

2.3 Procedure

Children were tested individually in a quiet area of their school on all four tasks. Each experimental session lasted approximately 20 to 30 minutes. Children were told that they were going to play some fun games with the experimenter and were rewarded with stickers throughout the experimental session. The metaphor task was always administered first, followed by the alternative naming task. The analogy perception task was administered at the end of the grid game session to limit as much as possible the effect of change in the procedure. The naming-and-pointing game for vocabulary assessment was administered at the end of the experimental session to prevent it from affecting answers to the metaphor task.

In what follows, the general procedure of the grid game is described first, followed by details on the procedure for every single task.

2.3.1 The Grid Game

The general procedure of the Grid Game was adapted from Morisseau, Davies & Matthews (2013). Children were presented with 3 boxes placed close to the grid. Each box contained the set of pictures for each of the experimental tasks. Children were told that the game was to arrange the pictures in each box according to a configuration that they could not see and that the experimenter would give them instructions to achieve this. The correct configuration was pictured on a booklet that the children could look at only at the end of each task. For every task there were 2 pictures configurations, one for the warm-up phase and the other for the test phase (see Figure 4).

Each of the experimental tasks was composed of a warm up, a test phase and a tidy-up phase. During warm up, children were asked to put on the grid all the items contained in the box according to the experimenter's instructions. The aim was to arrange the pictures according to the first configuration on the booklet (Configuration 1). This procedure was aimed at presenting and familiarizing children with the experimental items and procedure.

In the test phase, the aim of the game was to arrange the whole grid in order to match the final Configuration 2. Children were told that, in order to do so, some pictures had to be removed from the grid and placed into a basket. Thus, the experimenter would tell them which picture had to be removed and their task was to detach the pictures referred to and put them into the basket. Depending on the task, the target items were referred to in such a way as to elicit the use of one of the cognitive abilities under scrutiny. For the metaphor and alternative naming tasks, the procedure was exactly the same (i.e., moving the pictures in the grid to match a final configuration). What changed was the way in which the experimenter referred to the target picture, using a metaphorical or literal label in the metaphor task and the same or different label in the alternative naming task. Using the same grid game with the same attaching/detaching procedure, the task for analogy was slightly different because of the different nature of the phenomenon at stake: this time children were asked to choose a picture which best completed the presented sequence.

The tidy-up phase was used for the transition to the next task. Children were asked to help the experimenter remove all the pictures from the grid in order to open the next box corresponding to the next experimental task (i.e., box 1: metaphor task; box 2: alternative naming task; box 3: analogy task).

For all tasks, during both warm up and test phase, while giving instructions to the child, the experimenter always looked at the configuration on the booklet and never looked at the pictures in the grid or at the child. This was done to prevent children from making their choice on the basis of any non-verbal cue.

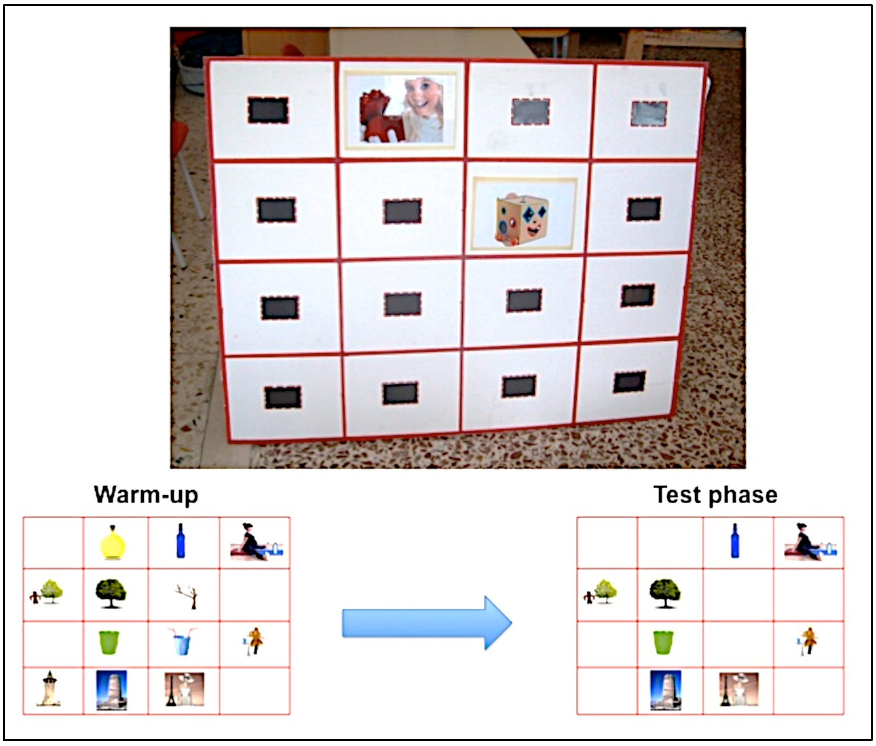


Figure 4: The Grid Game and two sample pictures for Configuration 1 (i.e., Warm-up) and Configuration 2 (i.e., Test phase) used in the metaphor task.

2.3.1.1 The Metaphor Task

The procedure for the metaphor task was adapted from Deamer & Pouscoulous (submitted). Children were asked to take pictures off the grid. These were referred to by using a metaphor in the metaphoric condition and a literal expression in the literal condition (e.g., *take the glass with antennae* and *take the glass with the straws* respectively). In both conditions, expressions of the form "The X with the Y" were used.

The metaphor condition (N. 4 trials) was always administered after the literal condition (N. 4 trials) to avoid any spill-over effect. For both conditions, in the warm-up phase, the experimenter first introduced the item to the child and then told her to put it in a specific slot of the grid (pretending to follow Configuration 1). All items of the triplet (i.e., target and controls) were presented to the child by using general and literal expressions referring to the target of the metaphor and never to its vehicle (e.g., Item 1: *Look! What a wonderful glass!;* Item 2: *Look! One more glass!;* Item 3: *What do we have here? Another* *glass!* for the *glass with the antennae* trial). This prevented any additional form of alternative naming to affect the metaphor task. In the literal condition, naming occurred following the same criteria as in the metaphoric condition and no reference was made to the target property (e.g., Item 1: *Look! What a wonderful glass!;* Item 2: *Look! One more glass!;* Item 3: *What do we have here? Another glass!* for the *glass with the straws* trial). For each trial, the three items were always presented consecutively and were always placed on contiguous slots of the grid to make sure that - in the test phase - children would pay attention to all 3 alternatives. For each triplet, the order of items presentation was randomized.

Once Configuration 1 was completed, the experimenter introduced Configuration 2 (i.e., test phase) by saying that the grid had to match a new picture but that, this time, there were too many pictures on the grid and that some of them had to be removed in order for the grid to match the new configuration. In the test phase, the experimenter told the child which picture had to be removed and referred to the target object using a metaphor (i.e., *take the glass with the antennae*) or a literal expression (i.e., *take the glass with the straws*). When the child did not choose, the experimenter repeated the instruction one more time. After the second instruction, if the child still failed to make the choice, then the experimenter named the target object literally (e.g., *Oh, I was distracted, I meant take the glass with the straws*).

2.3.1.2 The Alternative Naming Task

The procedure of the alternative naming task was adapted from Matthews, Lieven & Tomasello (2010). The target pictures were referred to by using the same label in the Same Term condition (S.T.) and 2 different labels in the New Term condition (N.T.). The S.T. condition was always administered before the N.T. condition to avoid spill-over effects.

During warm-up, all items were introduced to the child by referring to the depicted object 3 times (e.g., 1: *Look! A truck!;* 2: *I really like this truck*; 3: *Do you like this truck?* for a picture depicting a truck). This was done to make children memorize Label 1 and elicit their sensitivity to the use of a second label in the N.T. condition. In the test phase, the experimenter instructed the child about the pictures which had to be removed from the grid. In the S.T. condition, the target object was referred to by using the same label as used in the warm-up phase (e.g., *take the truck*). In the N.T. condition, the target object was referred to by using a different label from the one used in the warm up (e.g., *take the lollipop/candy*). If the child did not pick a picture, the instruction was repeated one more time. If the child still did not pick a picture, the experimenter named the target object using label 1 (e.g., *Oh, how headless I am! I meant take the lollipop*).

2.3.1.3 The Analogy Task

Analogy Perception was tested within the grid game scenario but, given the different nature of the ability at stake, the procedure slightly differed from the metaphor and alternative naming tasks. The procedure was a picture-sequencing task adapted from Goswami & Brown (1989, Experiment 1). The task was presented as a game about matching pictures and consisted of 4 trials. For each trial, all pictures (A & B sequence and 3 alternatives) were presented to the child and arranged in the grid according to the experimenter's instructions by naming the salient property of terms A and B. The pictures constituting the analogy sequence were introduced one at a time: first, picture A was introduced (*Look! This is a flying butterfly, put it in this slot of the grid*); second, picture B was introduced and the child was told to put it close to picture A because of the shared property (*Look! This is a bird and it flies too. So, put it close to the flying butterfly*). In order to create the sequence, pictures A and B were arranged on 2 contiguous slots of the grid. Then, the experimenter introduced to the child the triplets of alternative choices, which were arranged in contiguous slots of the grid but on a different line from the slots occupied by pictures A and B. In this phase, the experimenter asked the child to name the objects depicted in the pictures to ensure children recognized them.

To administer the task in a more implicit form, the analogical problem was presented to the child by induction. Testing analogy perception in a scenario where the analogy is more implicitly presented is crucial because in a metaphor the relation between target and vehicle is typically left implicit. Thus, for example, *the animals that fly* trial was phrased as follows: *if the flying butterfly is close to the flying bird, what do you put here* (pointing to the empty slot of the sequence) *to complete the sequence? This* (pointing at alternative 1), *this* (pointing at alternative 2) *or this* (pointing at alternative 3)? If the child did not choose among the three alternatives, the analogical problem was repeated one more time. If the child still failed to make a choice, then the experimenter solved the problem and explained it to the child (in which case, the trial was given accuracy 0).

After completion of the first trial, the pictures were removed from the grid and the next trial was introduced. The pictures of the preceding trials were removed because these distracted the child too much during piloting.

The slots of the grid used for terms A and B and for the alternative choices were the same across trials and participants. The order of presentation of the triplets of pictures was randomized across trials and participants.

2.3.2 Vocabulary assessment

The vocabulary used in the metaphor task was assessed at the end of the experimental session through a naming-and-pointing picture book adapted from Pouscoulous & Tomasello (2011). The comprehension part of the assessment came first and was followed by the production part. Comprehension was assessed by asking the child to point to the objects named by the experimenter (e.g., *Where are the antennae? Where is the straw?*). Production was assessed by asking the child to name the objects the experimenter pointed to (e.g., *What is this?).*

2.3.3 Data collection and coding

In both the metaphor and the alternative naming tasks, accuracy (i.e., correct selection of the target picture) and response times were collected. Accuracy was coded either as 1 or 0 depending on whether the child picked the right picture or not. The picture the child picked from the grid was considered her choice. Response times were recorded from when the experimenter uttered the experimental sentence to when the child provided her response and were computed by subtracting the child's response time to the time of the utterance.

For the analogy task, the child's accuracy was coded as 1 for the correct picture and 0 for the two incorrect pictures. The child's choice was considered made when the child attached her selected picture to the analogical sequence. Because of the nature of this task, which involved problem solving rather than the processing of a linguistic input, reaction times were not collected.

In the vocabulary assessment task, children were assigned one score per each correct answer. For comprehension, children's response was deemed correct when the child pointed to the target picture. For production, it was deemed correct when the child produced a literal label for the indicated object; correct literal labels included synonymous and/or dialectical expressions and words with slightly imprecise denotations, as long as these were not overextensions of the term (e.g., finger for hand). Following parental consent, all experimental sessions were filmed. All video-recordings were coded by the experimenter and a blind coder coded 20% of the data. Agreement was 100%.

3. Results

The proportion of correct responses and mean reaction times for the metaphor task per age group are reported in Table 3, together with the proportion of correct responses and mean reaction times in the alternative naming task and the accuracy rates relative to the analogy task. Results for each task are reported below in separate sub-sections.

As a general procedure, statistical analyses were carried out using Generalized Linear Mixed Models statistics (GLMMs) for accuracy and Linear Mixed Models (LMMs) for reaction times. The random structure of these models included random intercepts for subjects and items and by-subjects random slope for condition. Age Group, condition and their interaction were the fixed effects. LMMs statistics was conducted using the lme4 package (Bates et al., 2015) and the lmerTest package in the R environment (R Development Core Team, 2006) to extend the lmer model and provide the F statistics with degrees of freedom and p-value. GLMMs statistics was conducted using the glmer function from the lme4 package with a logit link and Laplace Approximation.

Whenever needed, a simple effects analysis of variance (i.e., for LMMs) and a Z-Test with continuity correction for the analysis of proportion in each group (i.e., accuracy data) were conducted, as well as Kendall's Tau correlations to better identify any developmental effect. The analyses of errors for the metaphor and analogy perception tasks were conducted with a chi-square statistics with Yate's continuity correction. Finally, to statistically assess whether Alternative Naming and Analogy Perception might significantly predict children's scores on metaphorical interpretation, a GLMMs statistics was carried out. Details of all results, including all values for non significant effects are reported in tables I to V at the end of the chapter.

Table 3. Proportion of correct responses (Mean Accuracy and Lower CI - Upper CI) and Reaction Times in seconds (mean and standard deviation) for all experimental tasks. Metaphor and Alternative Naming Tasks: Accuracy and Reaction Times; Analogy Task: Accuracy.

Metaphor Task				
	Accuracy		Reaction Times	
	Met	Lit	Met	Lit
3-year-olds	0.45 (0.34 - 0.57)	0.77 (0.67 - 0.87)	7.68 (6.74)	6.12 (3.92)
4-year-olds	0.56 (0.45 - 0.67)	0.93 (0.88 - 0.99)	4.72 (1.78)	4.52 (2.18)
Alternative Naming Task				
	Accuracy		Reaction Times	
	N.T	S.T	N.T	S.T
3-year-olds	0.77 (0.68 - 0.87)	0.98 (0.95 - 1)	7.13 (6.30)	4.06 (2.79)
4-year-olds	0.96 (0.92 - 1)	1 (1 - 1)	3.73 (2.59)	3.06 (2.41)
Analogy Perception Task				
	Accuracy		//	
3-year-olds	0.29 (0.18 - 0.40)		N.A.	
4-year-olds	0.40 (0.29 - 0.50)		N.A.	

3.1 Metaphor Results

Both in the metaphorical and literal conditions, children of both age groups always picked one of the three alternative pictures (i.e., Target, Target-Literal, Vehicle-Literal) for each trial. Both groups performed above chance level (i.e., 0.33) in both conditions as their probability of correctly selecting the correct (metaphoric) picture falls outside the 95% Wald CI of the estimated mean proportion of correct answers (Table 3).

The GLMMs statistics on accuracy data revealed that, compared to literal language, pre-schoolers of both age groups were significantly less accurate at correctly interpreting metaphorical expressions as shown by a significant main effect of Condition (χ^2 (1)=18.97; p<0.0001). This analysis also revealed a main effect of Age Group (χ^2 (1)=3.85; p<0.05), while the interaction ConditionXAge Group only approached significance (χ^2 (1)=2.90; p=0.08), presumably because literal language too is subjected to a developmental pattern in children of the targeted ages. To better characterize any developmental pattern in metaphorical competence, a two-sided Z-Test was conducted separately per age group and this confirmed that the proportion of correct choices in condition metaphor vs. literal significantly differed in both groups (3-year-olds: χ^2 (1)=13.37; p<0.0005; 4-year-olds: χ^2 (1)=28.03; p<0.0001). In addition, since literal language too is still developing in children of the targeted ages, a Kendall's tau correlational analysis was conducted on the accuracy of metaphoric trials only. This analysis revealed no Age Group effect (r_t= -0.10; p = n.s.).

An error analysis was conducted on metaphoric trials with accuracy 0 (see Figure 5). The chi-squared statistics showed that the most common error was control item *Vehicle-Literal* both in 3-year-olds ($\chi^2(1) = 23.21$; p<0.0001) and in 4-year-olds ($\chi^2(1) = 32.91$; p<0.0001), with no significant difference between age groups ($\chi^2(1) = 0.20$; p = n.s.). Thus, for example, when asked to give the experimenter *'the glass with the antennae'* (for the picture of a glass with two straws), children of both age groups tended to pick the picture depicting both a glass and a girl wearing an antennae-headband.

LMMs statistics on reaction times revealed a significant main effect of Age Group only (F(1, 35.64)=7.93; p<0.01) and no significant effects of Condition and GroupXCondition (all p_s = n.s.). To check whether these null results were possibly biased by delayed response times to incorrect trials (i.e., Accuracy = 0), a second LMMs analysis on reaction times of correct trials only (i.e., Accuracy = 1) was conducted. Apart from confirming the effect of Age Group (F(1, 31.69)=7.15; p< 0.05), this analysis revealed no significant effects of Condition and ConditionXAge Group (all p_s = n.s.).

Overall, these results suggest that both at age 3 and 4 children could interpret novel metaphors. Yet, children of both age groups exhibited more difficulties interpreting metaphorical than literal expressions and, when making a mistake, they interpreted the metaphor literally. Finally, interpreting metaphorical versus literal language does not seem to take longer both in 3- and 4-year-olds.

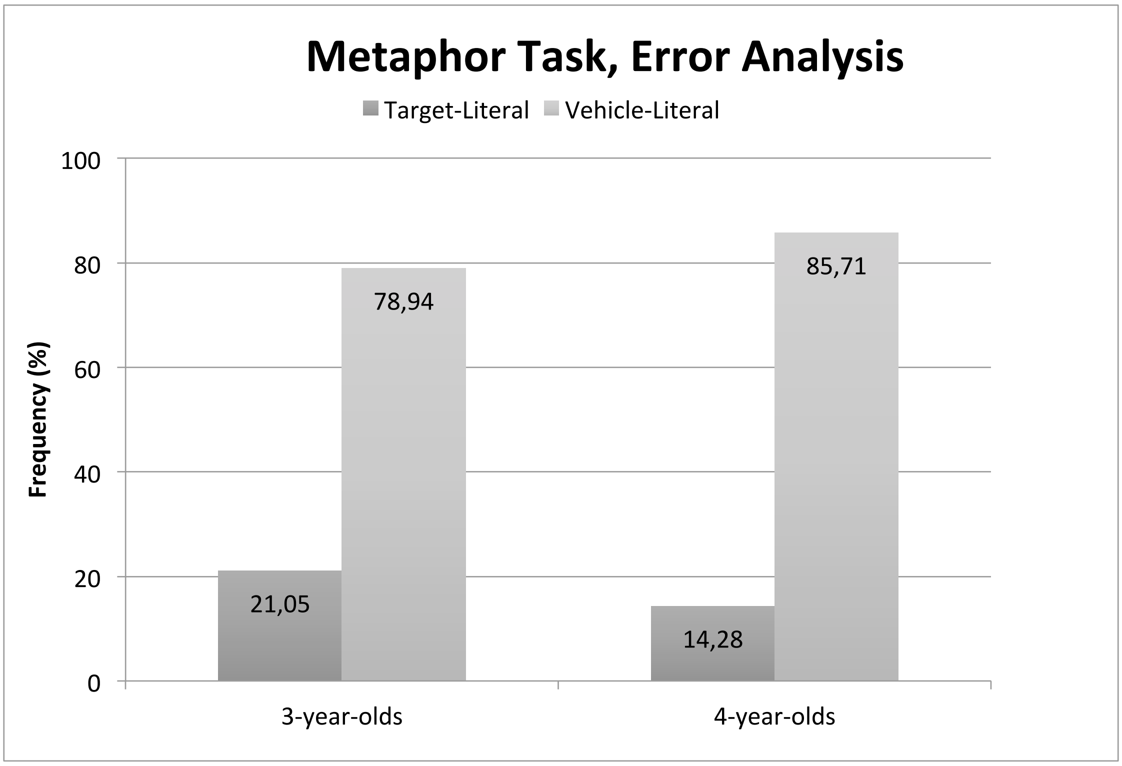


Figure 5: Frequency (%) of type of choice in metaphorical trials with accuracy = 0 per age group.

3.1.1 Vocabulary assessment: Results

Children of both age groups performed very well on the pointing-and-naming picture book task used for vocabulary assessment. Three-year-olds provided correct answers 88% of the time in the comprehension sub-task and 86% of the time in the production sub-part. The frequency in percentage of 4-year-olds' correct answers was slightly higher than 3-year-olds, with 91% correct answers both for comprehension and production. However, GLMMs statistics with Age Group as the between fixed factor revealed that the above difference was not significant for both comprehension ($\chi^2(1) = 1.73$; p=n.s.) and production ($\chi^2(1) = 1.45$; p=n.s.).

To exclude any effect of vocabulary on children's understanding of metaphors, children's comprehension and production scores were added as predictors to a GLMMs including metaphor accuracy as the dependent variable and condition as the independent variable[2]. This analysis revealed that children's accuracy in the metaphor task did not correlate with their knowledge of the vocabulary used (Comprehension: $\chi^2(1) = 0.80$; p=n.s.; Production: $\chi^2(1) = 0.04$; p=n.s.; ConditionXComprehension: $\chi^2(1) = 0.50$; p=n.s.; ConditionXProduction: $\chi^2(1) = 0.55$; p=n.s.). Kendall's tau correlations were conducted within each age group between children's accuracy in the metaphoric trials and their scores in both vocabulary comprehension and production tasks confirmed this pattern in both groups. Three-year-olds' performance in the metaphor task was not correlated with their mastery (i.e., comprehension and production) of the vocabulary used (Comprehension: r_t = - 0.07; p = n.s.; Production: r_t = -0.03; p = n.s.). The same pattern emerged with 4-year-olds: their ability to interpret metaphorical expressions was not correlated with their knowledge of the vocabulary used in the metaphor task (Comprehension: r_t = 0.02; p = n.s.; Production: r_t = -0.09; p = n.s.).

Overall, these results suggest that the vocabulary used in the metaphor task was age-appropriate and children's difficulty with metaphor interpretation was not ascribable to any lexical difficulty.

3.2 Alternative Naming: Results

Accuracy

In line with previous work on alternative naming (e.g., Matthews et al., 2010), overall both 3-and-4-year olds' performance on this task was pretty high and significantly above chance: the expected probability of picking the correct picture was 0.50, which falls outside of the 95% Wald CI of the estimated proportion mean of correct answers for both groups (Table 3). However, GLMMs statistics revealed a significant main effect of Condition ($\chi^2(1)$ = 9.75; p<0.0001) and Age Group (($\chi^2(1)$ = 1.42; p<0.0001), as well as a significant interaction

[2] Age group was not included as a fixed term in these models for two reasons. First, the statistics on vocabulary scores did not reveal any group-related effect for both comprehension and production. Second, when Age Group was inserted, the models did not converge (see Barr et al., 2013 for issues about convergence of GLMMs). However, since no group differences emerged in vocabulary assessment, keeping the Age Group out of the GLMMs fixed structure should still produce reliable results.

ConditionXAge Group ($\chi^2(1) = 14.55$; $p<0.0001$). The analysis of proportion of correct responses, conducted separately per age group, revealed that, contrary to 4-year-olds ($\chi^2(1) = 1.38$; p=n.s.), 3-year-olds were significantly less accurate at selecting the target picture when an alternative label was used ($\chi^2(1) = 13.07$; $p<0.0005$).

When not accepting the second label in the N.T. condition, 3-year-olds protested and corrected the experimenter (e.g., 'no, this is not a candy, it's a lollipop') 72% of the time; and they could not find the referred object 28% of the time. This occurred always in the N.T. condition and never with 4-year-olds.

Overall, when using an alternative label during linguistic reference, children picked the correct picture significantly above chance already at age 3 and this ability further improved as a function of age group, reaching ceiling levels by age 4. However, contrary to 4-year-olds, 3-year-olds still performed significantly better in the S.T. condition than in the N.T. condition, showing some difficulties at accepting an alternative label that mostly emerged by protesting or correcting the experimenter.

Reaction Times

As suggested by previous literature (e.g., Matthews et al., 2010), kids' sensitivity to alternative naming is known to particularly emerge in reaction times. Overall, the younger children were sensibly slower at performing the alternative naming task, as confirmed by the between-group analysis, which revealed a significant main effect of Age Group ($F(1, 36.77)=13.49$; $p < 0.001$). More importantly, this analysis revealed that the interaction ConditionXAge Group was significant too ($F(1, 43.78)=6.84$; $p < 0.01$).

To break down the above significant interaction, a simple effects analysis of variance was conducted where the effect of condition was computed across levels of the Age Group factor. This analysis confirmed that 3-year-olds were significantly slower in the N.T. than the S.T. condition ($F(1, 9.92)=5.22$; $p < 0.05$), while this was not the case for 4-year-olds ($F(1, 9.54)=0.22$; p=n.s.). In addition, Kendall's correlation conducted on the difference reaction times (N.T. minus S.T.) revealed that there was a significant negative correlation with Age Group ($r_t = - 0.21$; $p < 0.001$).

Taken together, these results suggest that children's difficulties associated to alternative naming lower as a function of age group: they are still visible at 3 and seem to be fully resolved by age 4.

3.3 Analogy Perception: Results

The probability of successful completion of the analogical sequences was relatively low and below chance level in both age groups: expected chance probability of selecting the correct picture was 0.33 which falls within the 95% Wald CI of the estimated proportion mean for both groups (Table 3). Kendall's correlation revealed no significant effect of Age Group (r_t = 0.10; p = n.s.).

Yet, an analysis of errors separately per age group and across groups revealed an interesting pattern of results (see Figure 6). The chi-squared statistics revealed that the effect of age group on the proportion of incorrect choices was significant ($\chi^2(1) = 4.38$; $p<0.05$). Interestingly, while in the younger group the difference of choice between the two control items was not significant ($\chi^2(1) = 0.04$; p = n.s.), in the older group it was highly significant indeed ($\chi^2(1) = 15.04$; $p < 0.0001$).

Overall, these results suggest that by age 4 analogical perception - as assessed in this experiment - is not yet developed enough to successfully complete a three-terms analogical sequence. However, a developmental pattern emerged from error analyses. While at 3 years of age children indiscriminately chose one of the 2 control pictures, thus probably selecting the preferred object, at age 4 it seems that children were following an analogical strategy: they could identify the salient property but they couldn't yet apply it to the correct object, as indicated by the fact that their preferred choice was control picture *Correct property/wrong object*.

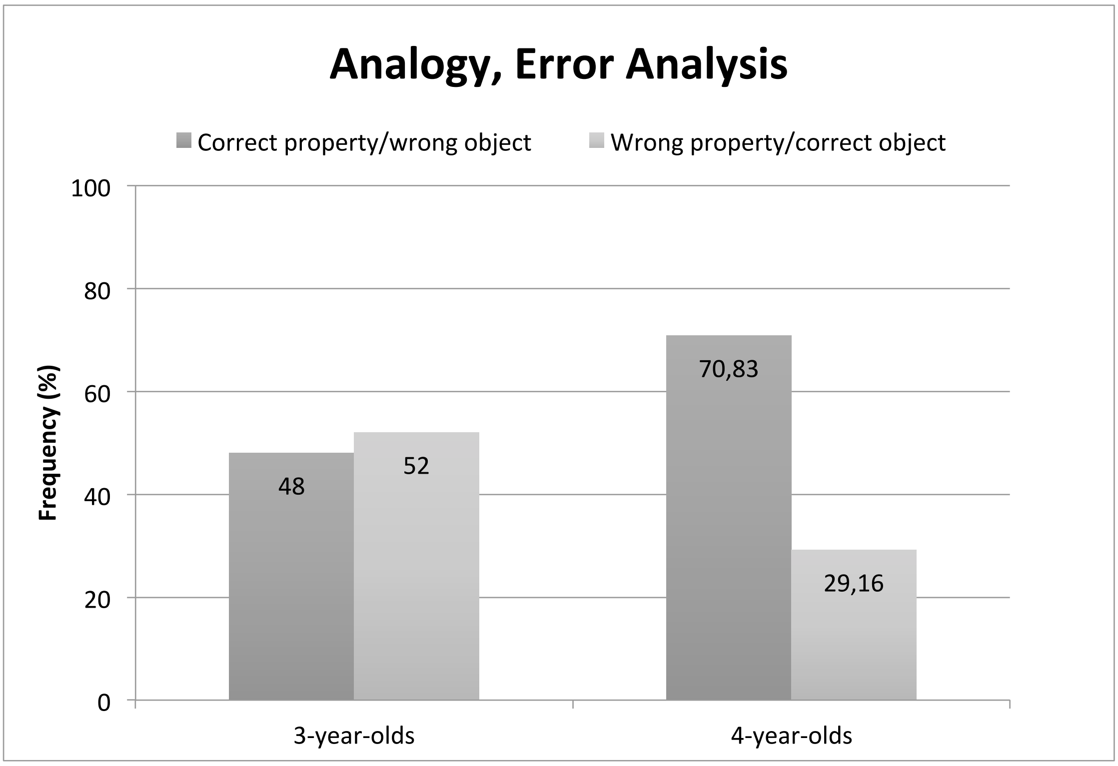


Figure 6: Frequency (%) of type of choice in analogical trials with accuracy = 0 per age group.

3.4 Multiple factors in metaphor development: Results

A GLMMs statistics was conducted with participants' accuracy in the metaphor task as the outcome variable, while Condition (i.e., literal vs. metaphorical) and participants' mean accuracy scores from the analogy and the alternative naming tasks as well as participants' averaged difference reaction times from the alternative naming task were the predictors. Within this model, Age Group (i.e., 3- vs. 4-year-olds) was not inserted as a predictor of metaphorical competence since (i) the results on the metaphor task revealed a main effect of age group for both the metaphor and literal condition; and (ii) the main aim of this analysis was to assess whether pre-schoolers' alternative naming and analogical abilities influence their metaphorical competence, independently of age group.

GLMMs results confirmed the significant main effect of Condition ($\chi^2(1) = 95.06$; $p<0.0001$) and revealed significant main effects of Alternative Naming (Accuracy: $\chi^2(1) = 988976$; $p<0.0001$; Difference RTs: $\chi^2(1) = 1326$; $p<0.0001$) and Analogy Perception ($\chi^2(1) = 314334$;

p<0.0001). Most importantly for the purpose of the present work, there were significant correlations between children's accuracy in the metaphor condition and each of the measures collected for the cognitive predictors at stake, as revealed by a significant interaction Metaphor ConditionXAnalogy ($\chi^2(1) = 153.518$; $p<0.0001$; $\beta=0.04$; SE=0.003347; z= 12.4); a significant interaction Metaphor ConditionXAlternative Naming Difference Reaction Times ($\chi^2(1) = 32.08$; $p<0.0001$; $\beta= - 0.01$; SE=0.003343; z= - 5.7) and a significant interaction Metaphor ConditionXAlternative Naming Accuracy (($\chi^2(1) = 503839$; $p<0.0001$; $\beta=- 2.37$; SE=0.003347; z= - 709.8). Furthermore, a likelihood-ratio test indicated that removing the above predictors from the model significantly decreased the goodness of fit ($\chi^2(7)= 56.28$; $p<0.0001$).

In summary, the GLMMs analysis revealed that pre-schoolers' alternative naming and analogical abilities significantly predict their metaphorical competence: the better children were in the analogy task and the quicker they were in the alternative naming task, the better their metaphorical competence (as indicated by the positive and negative correlations with Analogy and Alternative Naming RTs, respectively). Conversely, children who did better in the metaphor task, also did worse at accurately selecting the target picture in the alternative naming task (i.e., negative correlation between metaphor accuracy and alternative naming accuracy).

4. Discussion

4.1 Pre-schoolers' Metaphorical Competence

In the present experiment, two measures were collected in order to assess pre-schoolers' understanding of metaphors, namely reaction times and accuracy.

This study is the first collecting reaction times for metaphor understanding with very young children. Results on reaction times revealed no relevant effects (i.e., Condition and/or Age GroupXCondition), suggesting that both 3- and 4-year-olds were equally fast at picking a picture referred to metaphorically and literally. There are at least two main points worth noticing here. First, this null result is consistent with research on adults' comprehension revealing that the interpretation of metaphorical language does not necessarily elicit longer processing times than the interpretation of literal language (e.g., McElree & Nordlie, 1999; Glucksberg, 2001, 2003, 2008 and Gibbs, 1994). Overall, this fits well with theoretical models of parallel processing, claiming that metaphorical meanings are accessed directly without necessarily interpreting and rejecting the (incorrect) literal meaning first (Wilson & Carston, 2007; but see also Glucksberg, 2003). It is possible, then, that a similar interpretive process - not based on a rejection of the literal meaning of a metaphor - takes place in development too.

Second, an alternative explanation might be that no significantly delayed reaction times were observed with metaphor because literal language too is still developing in children this young and this might have contributed reducing any possible difference to a level which did not reach statistical significance. The only available study in which processing times were collected is Noveck et al. (2001). These results do not replicate Noveck et al., where instead increased processing times for metaphorical than literal expressions were found. However, substantial differences characterize Study 1 and Noveck et al. (2001): first, they tested much older children (i.e., from 8 to 14 years old) and, second, they collected the reading times of metaphorical and literal referents (e.g., toads vs. students) presented in the context of an eight-line story. Thus, there are neat developmental differences due to the different age ranges at stake between studies. Moreover, both the experimental task and material were extremely different too.

More interesting are the accuracy data. Three main results emerged from statistical analyses. First, all children were significantly less accurate at picking the correct picture when it was referred to metaphorically than literally (i.e., main effect of condition in GLMMs for accuracy). Second, when making mistakes, the most common error in both age groups was the *Vehicle-Literal* control, that is a picture where the two key referents of the metaphor (e.g., the glass and the antennae for "the glass with the antennae" trial) are shown literally (i.e., a picture depicting both a glass and a girl wearing an antennae headband). Third, children's performance in the metaphor task did not correlate with their mastery of the vocabulary used (i.e., vocabulary assessment analyses).

Compared to literal expressions, pre-schoolers' accuracy in selecting the target picture was significantly lower and this indicates that all children exhibited more difficulties interpreting metaphorical than literal language. To the best of my knowledge, this study is the first testing children this young with both metaphorical and literal expressions with a non-metaliguistic task. Contrary to previous recent studies not testing children in a literal condition (e.g., Pouscoulous & Tomasello, 2011; Deamer & Pouscoulous, submitted), it is possible that this manipulation made pre-schoolers' metaphorical difficulties emerge more clearly. At both age 3 and 4, accessing novel metaphors is a burden for children: unlike literal language, when a metaphor is used, they do not always understand the figurative expression and tend to assign a genuine literal interpretation by attending the literal meaning of both the metaphor target and vehicle, as the error analysis shows. Importantly, results on vocabulary assessment suggest that this is not ascribable to any lexical difficulties. In other words, it is not the case that children this young have difficulties accessing metaphorical meanings because they do not understand the meaning of each of the metaphor components: when a glass with two straws is metaphorically referred to as "the glass with the antennae", their ability to surf figurativity is somehow hindered, even though they do know what a glass is, what antennae are and what straws are. However, what said above does not mean that metaphorical competence is fully absent in pre-school years. In fact, children of both age groups performed above chance.

When compared to classical studies on the development of metaphor understanding (e.g., Nippold, 1988/1998; Winner, 1988/1997), overall, these results show that metaphoric competence develops earlier than adolescence - as classically argued - and provide further confirmation that more child-friendly experimental designs make children's metaphorical abilities emerge more clearly. The traditional view according to which metaphor understanding is achieved from age 8 and before this age children always interpret a metaphor literally is disconfirmed by these results. Rather, the emerged developmental pattern is in line with the more recent literature on metaphor development showing signs of metaphor understanding from age 3 and showing that children's vocabulary proficiency does not correlate with metaphor comprehension (e.g., Deamer & Pouscoulous, submitted; Pouscoulous & Tomasello, 2011). Taken together, this suggests that signs of metaphor understanding are already present in pre-school years: children are not always able to grasp the meaning of a metaphor, but neither they are unable at all.

Contrary to Deamer & Pouscoulous (submitted), no effect of age group with metaphor understanding was found in the present study. However, this may be due to differences in task demands, which were overall higher in this experiment than in Deamer & Pouscoulous. In Deamer & Pouscoulous, metaphors were presented within a short story context and it is well known that context eases metaphor interpretation (e.g., Pynte et al., 1996; Gildea & Glucksberg, 1983; Yang et al., 2013; Bambini et al., 2016a). In addition, they used nominal *X is a Y* expressions (e.g., *Alex is a monkey)* which are probably better accessed by young children because the identity statements might provide some cues to anchor the metaphorical referent to the target picture, thus, overall facilitating children's response choice. In the present experiment, metaphors were presented without any supporting context and children saw one target (metaphorical) picture together with two literal competitors that provided no cues for inferring metaphorical meaning (e.g., no identity statements; no totally unrelated picture), other than child-friendly task and vocabulary. In this respect, the metaphor task of Study 1 did not smooth the cognitive load required for metaphor understanding since it mainly aimed at investigating what makes metaphor comprehension harder for pre-schoolers. This might have made the task more stringent even for 4-year-olds, thus reducing the developmental progression.

To summarize, Study 1 replicates the developmental pattern for metaphor interpretation as emerged in recent literature: signs of metaphorical competence are visible already in pre-school years. Despite this, children are not yet proficient comprehenders of metaphor. At age 3 and 4, they experience clear and greater difficulties with metaphorical than literal meanings and these difficulties are driven by factors that do not include the mastery of vocabulary.

4.2 Alternative Naming

In the alternative naming task, children were asked to pick the picture of an object referred to using a newly introduced term (e.g., Candy) with respect to a previously used one (e.g., Lollipop). The new term was always a meaningful and literal though less conventional word for the target object and, strictly speaking, it was the word for a different entity from the target object (i.e., a candy vs. a lollipop) but which could be felicitously applied to the target object. This allowed us to test the alternative naming mechanism in a form as much similar as possible to that involved in metaphorical language (i.e., the metaphorical label is always meaningful and always refers to a different entity from the actual discourse referent) without involving the use of metaphor and category violation. In order to successfully perform this task, children were required to accept that one and the same referent can take two linguistic labels in the same communicative circumstance, which implies not just the use of two words for the same thing, but (i) an understanding that something can be described in different ways, and (ii) an understanding that speakers can take alternative perspectives on one and the same entity (see Doherty et al., 1998; Perner et al., 2002, 2003).

The costs associated to this kind of alternative naming were measured by collecting children's reaction times and accuracy. Three main results emerged. First, the alternative naming ability improves as a function of age group. Second, at age 3 children exhibit difficulties which clearly emerge in terms of reaction times and residually linger in the frequency of correct responses: they are significantly slower in the N.T. than in the S.T. condition and, though being well above chance, still they are less accurate in the N.T. Third, these difficulties are fully solved by age 4 (i.e., no significant differences for 4-year-olds both in reaction times and accuracy data). This pattern of results fully replicate Matthews et al. (2010) – which this task was adapted from. As for the accuracy data, Matthews et al. did not conduct statistical analyses but they report high accuracy rates for both age groups and a trend in protesting/correcting the experimenter when the alternative naming is not accepted by participants, both of which seem to match these findings.

Children's performance in the N.T. condition, as emerged in this study and above all for 3-year-olds, is higher than the one reported by Doherty & Perner (1998), Perner et al. (2002) and Perner et al. (2003), but importantly the general developmental pattern is roughly the same: 3-year-olds had more difficulties than 4-year-olds and there was an effect of age group. The differences in children's performance between those studies and Study 1 might be due to differences in the tasks and related demands. In these studies, the authors conducted a series of experiments where children' s alternative naming ability was assessed with respect to synonyms (e.g., rabbit/bunny) and superordinate categories (e.g., rabbit/animal). The children were asked to judge a puppet's naming as correct (if the puppet produced the synonym) or incorrect (if the puppet produced the same term as the child). In the production version, children were asked to produce a different name from the one used by the puppet (e.g., say *rabbit* if the puppet utters *bunny* and vice versa). This kind of alternative naming task "forces children to […] exactly […] acknowledge that something can be a rabbit and a bunny" (Perner et al., 2003: 372) and requires the child to explicitly think about a metalinguistic judgement that also involves (i) a certain level of theory of mind (i.e., the child has to judge the *puppet*'s linguistic behaviour, in turn pretended by the *experimenter*), and (ii) much more developed inhibitory skills than those typically shown by children at age 3 (e.g., Carlson & Moses, 2001) since the child had to inhibit the puppet's response and/or her preferred one. Beyond this, the production version required the child to *produce* the synonym and it is known that comprehension develops earlier than production (e.g., Clark, 2009). Conversely, the task in the present study was much simpler: the child only had to detect the alternative naming (e.g., no judging, no production) and react to this by selecting a picture. As a consequence, then, the demands imposed by this alternative naming task were much smaller and this might explain Study 1 participants' higher performance.

However, overall, the results emerged in this study replicate the developmental pattern associated to alternative naming as depicted by both Matthews et al. (2010) and the series of studies by Perner and colleagues: alternative naming develops between age 3 – when, as a whole, children still exhibit some difficulties – and age 4, when their difficulties are fully overcome. This is in line with Perner's hypothesis that alternative naming abilities develop by age 4, when kids have sufficiently developed theory of mind functions which allow them to entertain mental representations where what they know is more fluidly integrated to the speaker's intention.

As for 3-year-olds, results revealed that they were significantly slower and less accurate when a new term was used compared to when the same previously used term was employed. Importantly, however, their accuracy rates in the N.T. condition were not only significantly above chance but reached on average 77% correct responses. This is interesting indeed since it shows that, under reduced task demands, 3-year-olds are able to perform an alternative naming task. They seem to understand what the referent for the alternative naming act is, but still have difficulties *accepting* it. In other words, they do understand that 'the candy' is actually the lollipop, they acknowledge it but do not always accept it. And, importantly, when they do, they are however slower than in the S.T. condition. Thus, overall, alternative naming is somehow taxing for 3-year-olds. Following Perner's work (see also Clark, 1997 and Tomasello, 1999 for compatible explanations), such extra-effort can be interpreted as evidence in support of the idea that 3-year-olds are more sensitive to the change in the reference act precisely because of their difficulties with alternative perspective, especially when this occurs within the same communicative/referential context. This is also consistent with Matthews et al. (2010) view that increased sensitivity to alternative naming is caused by the introduction of a change in perspective which, in turn, interferes with a previously constructed mental representation that selectively encoded what someone has said to them and how. According to Matthews et al., this violates kids' expectation about the referential pact (not manipulated but involved in the paradigm of the present study) and, when it happens, they tend to actively react by correcting the experimenter or protesting. All behaviours, these latter, which received accuracy 0 in this task since they all constituted a manifestation of difficulties at accepting the second label and prevented the children to successfully pick the target object.

4.3 Analogy Perception

The analogy task was aimed at assessing the extent to which analogy perception is developed in pre-school children to allow them to cope with metaphor comprehension. In order to comprehend a metaphor, it is necessary to access the metaphor ground, that is figuring out the relation between target and vehicle. Thus, at least two operations are needed: first, identifying a relevant property of the vehicle of the metaphor to, second, meaningfully apply it to the target. Importantly, identification and application of the relevant property have to be done in a context where the relation between vehicle and target is typically left implicit. All this involves a certain level of complexity which goes beyond basic analogical skills. Crucially, then, in the analogy perception task, maintaining such a level of complexity was of primary interest. This was done in two ways. First, to successfully complete the pictorial pattern with the target picture (e.g., the flying bee), children had not only to identify the relevant feature (e.g., flying, not sleeping), but also to apply it to the relevant object (e.g., the bee, not the helicopter). Second, the analogical sequence was presented by induction, without providing the child with any supporting cue which could drive analogical reasoning. This allowed us to test analogy perception in a more implicit form.

Both 3-and-4-year-olds completed the task by choosing at chance the target picture. These results indicate that children exhibited remarkable difficulties in a task where more than basic analogical abilities were required. In turn, this suggests that by age 4 children's analogy perception is not developed enough yet to figure out a relation which is only implicitly conveyed.

These findings are only partially consistent with Goswami & Brown (1989): in their experiment 1, 3-and-4-year-olds exhibited greater difficulties than 6-year-olds (who were at ceiling), which indicates that in pre-school years analogical abilities are still developing. However, unlike in the present study, the younger children were above chance. This difference can be interpreted in terms of task complexity, which was overall reduced in Goswami and Brown in several respects. First, in Goswami and Brown, children's highest performance was found in the explanation condition, where the experimenter always explained the child what the correct choice was and why. Thus, children showed better analogical abilities when the analogy was explained to them presumably because the reasoning was made explicit at the end of each trial and, over trials, this might have trained the child to reason analogically (see Tunteler & Resing, 2002, 2007 for effects of training in children's spontaneous analogical reasoning). In fact, when no explanation was provided (i.e., induction condition), children's performance dropped significantly. In addition, in the induction condition, even if no explanation was provided, children always received a feedback on the correctness of their choice and, when making mistakes, they were shown the right picture. This, in turn, might have prompted them to figure out the key mechanism for the patterns to be completed. In general, then, in Goswami and Brown, there were always some supporting cues which prompted the kids to reason by analogy. Conversely, in the present experiment, analogy perception was elicited at a level of complexity such that the analogy was inducted (i.e., more implicit form) and there were no supporting hints to facilitate the task.

The present findings - and the discussion above - are consistent with more recent work on the development of analogical reasoning showing that, in pre-school years, it is constrained by task complexity: children's performance significantly lowers when relational complexity and featural distractors are increased (e.g., Richland et al., 2006) and when the analogical transfer is unaided and untrained in spontaneous elicitation scenarios (e.g., Tunteler & Resting, 2002, 2007). It has been suggested that constraints in children's analogical reasoning and developmental changes are due to inhibitory control and working memory limitations, respectively needed to integrate multiple relations and inhibit irrelevant/distracting properties/entities (see Richland et al., 2006 for a review and Morrison et al., 2011).

As for the analysis of errors, while 3-year-olds indiscriminately chose one of the two controls, 4-year-olds chose the control item *Correct property/wrong object* (e.g., the helicopter) significantly more often than the control item *Wrong property/correct object* (e.g., the sleeping bee). These data are in line with Goswami and Brown's (1989) analysis of errors and, importantly, suggest that there is a developmental pattern. Unlike 3-year-olds, 4-year-olds were following a strategy that can be considered analogical since it was driven by the (successful) search for a relevant property. Hence, by age 4, children can identify a shared property (i.e., the helicopter flies too). What seems to be still developing is their ability to apply that relevant property to the relevant object, or - at least - to do so within a more complex scenario[3]. However, since both steps are needed to derive the (implicit) metaphor ground, this may in turn increase the cognitive demands associated to metaphor understanding and may contribute to slow down the developmental trend of metaphorical competence.

5. General Discussion: Metaphor development and multiple factors

The main goal of this study was to extend Rubio-Fernández & Grassman's (2016) findings and have an explanatory look at the development of metaphorical competence by looking at the role of alternative naming and analogy perception, two cognitive abilities involved in metaphor interpretation and still developing in pre-school years. The main hypothesis was that the above abilities might influence the development of metaphorical understanding and, considering their developmental trajectories, act as impeding factors likely to hinder pre-schoolers' comprehension of a metaphor. It was predicted that the more alternative naming and analogical skills are developed, the more metaphorical competence is

[3] Note that the fact that children chose the target item at chance in this experiment does not actually suggest a total lack of analogy perception because all 3 items presupposed a certain degree of analogical abilities. In fact, the two control items were adapted from Goswami & Brown (1989, items E & F), who used those items to check whether, even though making mistakes, children were still attempting to reason by analogy. Based on this, chance performance in the present study only suggests that, when elicited at a certain level of complexity, pre-schoolers' analogical abilities are not developed enough to simultaneously identify a shared property and apply it to the correct object within a context in which analogy perception is presented in a more implicit form.

developed too. Conversely, the less one or both of them is developed, the less metaphorical competence is developed. The results confirm this general prediction: overall, metaphor proficiency was found to significantly correlate with pre-schoolers' proficiency at accepting two labels for the same referent and drawing inducted (i.e., more implicit) analogies.

With respect to alternative naming, significant negative correlations with children's performances in the metaphorical task emerged. On the one hand, the quicker children were at performing the alternative naming task (i.e., reaction times), the better they were at interpreting metaphor. On the other hand, the better (i.e., more accurate) children were in alternative naming, the worse (i.e., less accurate) they were in the metaphor task. These two correlations are discussed separately below. However, a greater focus is given to data from reaction times for 3 main reasons: (i) they are the standard measure by which children's sensitivity to alternative naming is known to emerge in a referential paradigm (see Matthews et al, 2010); (ii) despite the significant age group differences, overall, children's accuracy rates in this task were very high and this suggests that pre-schoolers are able to deal with alternative labels pretty well already at age 3; (iii) rather than reflecting whether or not children understand and accept two linguistic labels for the same referent, reaction times reflect the time needed to process the phenomenon at stake and, as such, they provide more direct evidence on the underlying cognitive costs.

Let us start from reaction times. I found that the shorter the difference reaction times at alternative naming (i.e., between S.T. and N.T. conditions), the higher children's accuracy in the metaphor task. In other words, as long as accepting two labels for the same referent becomes easy enough not to require additional processing costs, interpreting a metaphor becomes easier too. I suggest this is so because a sufficiently developed alternative naming ability enhances metaphor interpretation since it provides the cognitive cornerstone needed to access and deal with metaphor dual reference. In addition, the developmental patterns emerged from the metaphor task on the one hand and the alternative naming task on the other hand strongly indicate that alternative naming is a source of difficulty for metaphor comprehension at age 3, but not at age 4: at age 3, the increased cognitive demands associated to alternative naming significantly tax the cognitive load underlying the comprehension of metaphorical language. For this reason, alternative naming can be reasonably considered an impeding factor for young kids' metaphor interpretation. This is very likely to be so until age 4, when their adult-like alternative naming abilities might reduce the cognitive load of metaphor interpretation. Apart from the previously cited literature on alternative naming, converging evidence in support of such a developmental hypothesis comes from Rubio-Fernández & Grassmann (2016). In this study, the authors found that, contrary to 3-year-olds, 4-year-olds performed above chance level in the condition requiring them to use alternative naming together with spatial analogy. This further strengthens the idea that, by age 4, children's processing resources are developed enough to deal with alternative naming.

The negative correlation between alternative naming accuracy and metaphor accuracy is less straightforward: children who were better at interpreting metaphors (i.e., more accurate), were also worse (i.e., less accurate) at accepting two labels for the same referent. This might be surprising at first since, similarly to reaction times, one would expect that accurate responses in alternative naming should lead to more accurate responses in a metaphor task. One possible explanation for this is that such a dissociation might actually reflect two different underlying processes. Accuracy in the metaphor task genuinely indicates children's understanding of a metaphor: in this experiment, if the child did not grasp the meaning of the metaphor, she simply could not select the target picture. In other words, there was no reason for the child to select the target picture, especially considering that she could opt for other (more or less suitable) referents provided by the control items. In fact, this is exactly what happened. As a consequence, then, accuracy in the metaphor task reflects pure understanding of the metaphor. However, this might not necessarily be the case for accuracy in the alternative naming task. Here, a child's response was given accuracy 0 in two main cases: either (i) the child could not find the referred object (28% of the time for 3-year-olds), or (ii) she protested and corrected the experimenter (e.g., "no, this is not a candy, it's a lollipop"), which occurred 72 % of the time with 3-year-olds. The case in (ii) was coded accuracy 0 because this prevented the child to finally pick the target object and was hence interpreted as an indication for the child's difficulty with alternative naming. Importantly, however, protesting at something or correcting the experimenter presupposes an understanding of which was the referred object and only reveals that children have more difficulties *accepting* an alternative label (see corresponding discussion section), but it does not exclude understanding. To put it simple, children did understand that the referred candy was the lollipop, but they were reticent accepting this. Note that the same does not apply to reaction times data since they are a genuine measure of the underlying cognitive costs for alternative naming. In fact, the correlation between reaction times data and metaphor accuracy indicated that the easier alternative naming (i.e., smaller reaction times), the easier metaphor understanding too.

With respect to analogy perception, a significant positive correlation emerged showing that children with more developed analogical abilities exhibit a better understanding of metaphor. Overall, this suggests that better developed analogical skills enhance the interpretation of metaphorical language in pre-school children, presumably because this prompts the identification of the properties and relations shared by the metaphor target and vehicle, that is it facilitates the recognition of the metaphor ground (see Rubio-Fernandéz & Grassmann, 2016 for a similar link between metaphor and analogy). However, results from the analogy task revealed that both 3- and 4-year-olds successfully completed the analogical sequences below chance level. Error analysis suggested that their analogical abilities were not developed to the same extent, though: contrary to 3-year-olds, 4-year-olds chose control item *Correct property/wrong object* significantly more often than control item *Wrong property/correct object*. This suggests that by age 4 analogy perception (as elicited in this experiment) is developed enough to enhance the identification of the relevant property shared by two or more entities, but it is not yet mature enough to perform some more operations on the property to be applied to the correct object. If, on the one hand this may improve 4-year-olds' metaphorical skills, on the other hand it can also explain their difficulties with metaphor understanding: in order to fully grasp the metaphor ground, it is necessary both to identify the vehicle relevant property and to correctly apply it to the metaphor target. Overall, then, at age 3 analogical perception seems not to be developed enough to this purpose and is likely to impede metaphor understanding. At age 4, children exhibit more developed analogical abilities that might allow them to somewhat access the metaphor ground. However, their analogical abilities are not yet developed enough to apply the identified property to the correct object and this might contribute to explain where their still observable metaphorical difficulties stem from at that stage of development.

A final point worth noticing is that in the alternative naming and the analogy perception task there was a stark improvement at age 4. In the analogy task, though making mistakes, 4-year-olds identified the relevant property. In the alternative naming task, children's difficulties were totally overcome by age 4. In general, this is worth noting since the progress from 3 to 4 in these tasks may be due to the development of inhibitory control, which was an underlying requirement in both tasks and which is known to pick precisely between age 3 and 4 (Carlson & Moses, 2001).

To summarize, pre-schoolers' metaphorical proficiency was found to be significantly predicted by alternative naming and analogy perception. The emerged findings show that the development of metaphorical competence is influenced by - and bound to - the development of alternative naming and analogy perception. At age 3, both cognitive skills seem to hinder metaphor interpretation. At age 4, alternative naming is not taxing anymore for children, while some costs are still associated to analogy perception: thus, while alternative naming should not represent anymore an impeding factor, analogy perception - at a level of complexity as involved in metaphor interpretation (e.g., implicit) - is likely to prevent children from genuinely appreciating metaphors.

This study is the first investigating directly the role of alternative naming and analogy perception in the development of metaphoric competence within a single experimental paradigm. The results fit well with the growing body of research currently indicating that the development of metaphorical competence seems to depend on a cluster of cognitive abilities (Pouscoulous, 2014; Rubio-Fernández & Grassmann, 2016; Deamer & Pouscoulous, submitted). This study showed that alternative naming and analogy perception are two such candidate skills. However, they are not the only ones responsible for the interpretive demands underlying children's metaphor comprehension. The most important canditate skill might be inhibitory control since metaphor requires the irrelevant properties of the literal meaning of the vehicle to be suppressed. In the adult literature, to the best of my knowledge, there are no studies directly investigating the role of ihnibitory control in metaphor interpretation, but it has been shown that literal meaning is always activated in early metaphor processing and suppressed later on (Rubio-Fernández, 2007) and that irrelevant information in metaphor comprehension is inhibited (Glucksberg et al., 2001).

Inhibitory control can explain the difference between children' s metaphorical abilities at age 3 and 4 that is observed in the literature since it picks in fact between 3 and 4 (e.g., Carlson & Moses, 2001; see Petersen et al., 2016 for a review). Furthermore, there is preliminary evidence that pre-schoolers' inhibition skills predict their metaphor understanding (Deamer & Pouscoulous, submitted).

One more candidate skill might also be working memory capacity, which is still very limited in early childhood (e.g., Simmering, 2012; Thomason et al., 2009; see Simmering & Perone, 2013 for a review). When interpreting a metaphor, not only there are several information to be actively maintained and updated in working memory, but also its cognitive complexity may rely in the simultaneous involvement of multiple factors (e.g., analogy perception, alternative naming, inhibitory control, etc.), which calls for a massive involvement of executive functions in general. There is consistent evidence that working memory correlates with adults' metaphor processing (e.g., Prat et al., 2012; Olkoniemi et al., 2016; Chiappe & Chiappe, 2007). As for research in development, it has been shown that executive functions play a central role in metaphor comprehension during adolescence and that these are restricted to updating information in working memory and cognitive inhibition (Carriedo et al., 2016). Nothing is known yet with respect to early childhood, but - for example - Rubio-Fernández and Grassmann (2016) might provide indications to this purpose as they found that 4-year-olds, but not 3-year-olds, were able to integrate analogy perception and alternative naming.

Finally, the results of the present study support well the account on language interpretation framed by Relevance Theory (Sperber & Wilson, 1986/95; Carston, 2002). According to relevance theorists, language comprehension is guided by an inferential pragmatic process based on the notion of *Relevance*. Relevance is a property of the inputs to cognitive processes, it is a function of effects and effort: the more the cognitive effects achieved, the more the relevance of that input; the more the cognitive effort required to process an input, the less the relevance of that input. In turn, the comprehension procedure to be followed is one in which the interpretation has to comply with a certain threshold of relevance (i.e., effects vs. effort) and the speaker's intended meaning is fully derived when expectations of optimal relevance are satisfied, that is when the cognitive effects achieved by an utterance outweigh the cognitive effort required to achieve those cognitive effects. It follows from this that utterance interpretation always carries some worthwhile cognitive effects (i.e., the derivation of the speaker's meaning), but all this comes at a cost (e.g., in terms of processing resources).

Overall, this study provides a further contribution to a relevance-oriented framework of language comprehension. In particular, it might be the case that pre-schoolers miss access to the effects of metaphor because the associated costs - required to achieve precisely those effects (cf. the notion of relevance) - are still too high in that phase of development, although they are not the same at age 3 and 4 (see also Noveck et al., 2001). The present study has showed that, in pre-school years, two of the factors which increase the cognitive costs underlying metaphorical language are alternative naming and analogy perception, respectively required to deal with dual reference and the metaphor ground. If one adopts a relevance theoretic perspective on language comprehension, then the above factors can be said to hinder the cognitive effects associated to metaphorical language.

To conclude, further research is clearly needed to better characterize the cluster of cognitive abilities involved in the development of metaphoric competence, to compare this with other apparently simpler figurative expressions and to replicate the findings. Currently, we know a lot about metaphor and we know a lot about language development, but our knowledge of what's in a metaphor is much deeper than our knowledge of what are the cognitive prerequisites for a developing mind to overcome the cognitive costs brought about by metaphor and successfully jump into metaphoricity, hence benefiting from its cognitive effects. The present study suggests that the cognitive recipe for the development of metaphoric competence includes alternative naming and analogy perception.

Table I - Results of all statistical analyses conducted on the Accuracy and Reaction Times data for the Metaphor Task.

Metaphor Task						
Accuracy						
GLMMs						
	β	SE	z	χ^2	DF	$P(\chi^2)$
Condition	-2.84	0.6743	-4.22	18.97	1	<0.0001***
Age Group	-1.68	0.6639	-2.53	*3.85*	1	<0.05*
ConditionXAge Group	1.18	0.6971	1.70	*2.90*	1	0.08 .
Analysis of Proportion per age group: two-sided Z-Tests with continuity correction						
3-year-olds						
Condition	//	//	//	13.37	1	<0.0005**
4-year-olds						
Condition	//	//	//	28.03	1	<0.0001***
Kendall's tau correlation on metaphor trials only						
	z	r_t	p			
Age Group	- 1.27	- 0.10	0.20			
Error Analysis: Chi-squared tests with Yate's continuity correction						
				χ^2	DF	$P(\chi^2)$
3-year-olds						
Control choice				23.21	1	<0.0001***
4-year-olds						
Control choice				32.91	1	<0.0001***
3- vs. 4-year-olds						
Age Group				0.20	1	0.65
Reaction Times						
LMMs						
	β	SE	t	F	DF	p
Condition	0.20	0.8146	0.25	1.59	1, 16.60	0.22
Age Group	1.60	0.7027	2.28	7.93	1, 35.64	<0.01*
ConditionXAge Group	1.35	0.8909	1.52	2.31	1, 55.86	0.13
LMMs on trials with Accuracy = 1 only						
Condition	0.39	0.6576	0.60	0.72	1, 12.05	0.41
Age Group	1.02	0.4642	2.21	7.15	1, 31.69	0.01*
ConditionXAge Group	0.21	0.6631	0.32	0.10	1, 35.54	0.74

Table II - Results of all statistical analyses conducted for Vocabulary Assessment.

Vocabulary Assessment						
GLMMs						
Comprehension						
	β	SE	z	χ^2	DF	$P(\chi^2)$
Age Group	-0.64	0.4938	-1.31	1.73	1	0.18
Production						
Age Group	-1.10	0.9165	-1.20	1.45	1	0.22
Analysis of Covariance: Vocabulary Score & Accuracy in Metaphor Task						
GLMMs, Comprehension						
Condition	-3.22	1.442	-2.23	18.95	1	<0.05*
Comprehension	-1.28	1.175	-1.09	0.80	1	0.36
ConditionXComprehension	1.003	1.410	0.71	0.50	1	0.47
GLMMs, Production						
Condition	-1.53	1.1326	-1.35	19.86	1	<0.0001***
Production	0.25	0.7850	0.32	0.04	1	0.82
ConditionXProduction	-0.83	1.1229	-0.74	0.55	1	0.45
Kendall's tau correlations per age group: accuracy in metaphor trials only & vocabulary						
	z	r_t	p			
3-year-olds						
Comprehension	- 0.63	- 0.07	0.52			
Production	- 0.25	- 0.03	0.79			
4-year-olds						
Comprehension	0.17	0.02	0.85			
Production	- 0.88	- 0.09	0.37			

Table III - Results of all statistical analyses conducted on the Accuracy and Reaction Times data for the Alternative Naming Task.

Alternative Naming Task						
Accuracy						
GLMMs						
	β	SE	z	χ^2	DF	$P(\chi^2)$
Condition	3.63	0.0028	125.794	9.75	1	<0.0001***
Age Group	-3.32	0.0027	-1.190	1.42	1	<0.0001***
ConditionXAge Group	-3.48	0.0028	-120.660	14.55	1	<0.0001***
Analysis of Proportion per age group: two-sided Z-Tests with continuity correction						
3-year-olds						
Condition	//	//	//	13.07	1	<0.0005***

4-year-olds						
Condition	//	//	//	1.38	1	0.23
Reaction Times						
LMMs						
	β	SE	t	F	DF	p
Condition	-0.61	1.2839	-0.47	2.26	1, 8.64	0.16
Age Group	3.40	0.9284	3.66	13.49	1, 36.87	0.0007**
ConditionXAge Group	-2.43	0.9327	-2.61	6.84	1, 43.77	<0.05*
Simple effects ANOVA						
3-year-olds						
Condition	-1.52	0.661	- 2.03	5.22	1, 9.92	<0.05*
4-year-olds						
Condition	- 0.30	0.642	- 0.47	0.22	1, 9.54	0.64
Kendall's tau correlation on difference reaction times (N.T - S.T.)						
	z	r_t	p			
Age Group	- 3.18	- 0.21	<0.001**			

Table IV - Results of all statistical analyses conducted on the Accuracy data for the Analogy Perception Task.

Analogy Perception Task			
Kendall's tau correlation			
	z	r_t	p
Age Group	1.32	0.09	0.18
Error Analysis: Chi-squared tests with Yate's continuity correction			
	χ^2	DF	$P(\chi^2)$
3-year-olds			
Control choice	0.04	1	0.84
4-year-olds			
Control choice	15.04	1	<0.0001***
3- vs. 4-year-olds			
Age Group	4.38	1	<0.05*

Table V - Results of the Generalized Linear Mixed Model statistics conducted for the analysis of Covariance between children's accuracy in metaphor task and their performance in each of the tasks for the predictors under scrutiny.

Predictors of Metaphor Comprehension						
GLMMs: Accuracy in Metaphor Task ~						
	β	SE	z	χ^2	DF	$P(\chi^2)$
Condition	0.032	0.003347	9.9	95.06	1	<0.0001***
Alternative Naming Accuracy	3.33	0.003348	994.7	988.976	1	<0.0001***
Alternative Naming Difference reaction times	0.12	0.003343	36.4	1.326	1	<0.0001***
Analogy Accuracy	1.87	0.003348	560.7	314.334	1	<0.0001***
ConditionXAlternative Naming Accuracy	-2.37	0.003347	-709.8	503.839	1	<0.0001***
ConditionXAlternative Naming Difference reaction times	-0.018	0.003343	-5.7	32.087	1	<0.0001***
ConditionXAnalogy Accuracy	0.041	0.003347	12.4	153.518	1	<0.0001***
Models Comparisons: Likelihood Ratio test Model 0: Met_Acc ~ 1 + (1+Subject) Model 1: Met_Acc ~ Cond + Predictors + Cond*Predictors + (1|Subject)						
	DF	AIC	BIC	χ^2	DF	P
Model 0	2	375.38	382.80			
Model 1	9	333.10	366.49	56.284	7	<0.0001***

Chapter II

Study 2 - Metaphor out of University: Electrophysiological correlates of Metaphor Processing in Academic and Non-Academic Learners

1. Introduction

A few years ago, a popular billboard on the London tube advertised the city of Las Vegas as *'a place where your accent is an aphrodisiac'*. The legendary rock'n roll star, Elvis Presley, opened one of his tracks singing *'You ain't nothin' but a hound dog'*. In other words, metaphor is ubiquitous in language use. It is a powerful communicative device to capture the audience's attention and marketing or music are only two such examples, many others being found for instance in politics and literature. Also everyday communication is densely populated by metaphors: 6 metaphors a minute, based on some estimations (H.R. Pollio, Barlow, Fine, Pollio, 1977). Everyone is repeatedly exposed to metaphor and experiences the process of understanding it.

Several theories and empirical works exist on the mental processes involved in metaphor comprehension. In particular, research has shown that the brain electrical correlates of metaphor processing differ from those associated with literal language interpretation (e.g., Rataj, 2014), and processing mechanisms vary as a function of the linguistic features of the stimulus itself (e.g., familiar vs. novel metaphors, single vs. extended metaphors, literary metaphors; e.g., Arzouan, Goldstein, Faust, 2007a; Bambini, Canal, Resta, Grimaldi, 2018; Lai, Curran, Menn, 2009; Rubio-Fernández, Cummings, Tian, 2016;). Much less is known on whether and how metaphor comprehension varies depending on socio-demographical factors such as individuals' education type and cognitive performance. A growing number of behavioural studies investigated variation in metaphor processing in relation to individuals' cognitive performance, especially working memory (e.g., Chiappe & Chiappe, 2007; Columbus et al., 2015), but the literature on the brain correlates is still limited. As for education, the effect of schooling has been documented for metaphor comprehension (Champagne-Lavau, Monetta, Moreau, 2012), and for the appreciation of literary texts (Burkett & Goldman, 2016; Peskin, 2010). However, as typical in the psychological literature (Norenzayan & Heine, 2005), the available data have been mostly collected from highly-educated university students. University students represent just the 31% of the world population (OECD Indicators, 2012), which raises some concerns for the generalization of results and for comprehensive theories of metaphor interpretation. The present study aims at moving a step forward in research on metaphor understanding, by investigating the impact of socio-demographical factors on metaphor processing, focusing in particular on one such factor: Education Type. The electrophysiological response of individuals with an academic and a non-academic education was registered while they were reading metaphorical and literal expressions. A battery of measures related to general cognitive performance, including working memory, was collected as well.

1.1 Metaphor Interpretation: traditional theories and recent developments

Metaphorical expressions are paradigmatic cases of language use where the sentence meaning underspecifies the speaker's intended meaning. For example, in a metaphor such as *Hope is a star*, the target (i.e., hope) is a concept described by means of another concept, the vehicle (i.e., star), typically belonging to a different semantic category and such that some properties of the vehicle are figuratively applied to the target. As a consequence, the speaker's intended meaning departs from what is literally said, that is from the semantic content of the utterance, which, in a metaphor, is always false (i.e., literally speaking, hope is not a luminous sphere of plasma).

Because of this blatant departure from literalness, the issue of whether or not metaphor interpretation involves accessing the (false) literal meaning is the cornerstone of the theoretical debate on metaphor (Bambini & Resta, 2012). As mentioned in the Introduction of this dissertation, within the pragmatic literature, two main traditional views have been developed. According to the indirect access accounts (e.g., Grice, 1975; Searle, 1979), metaphor is a special case of language use in which the speaker's intended meaning is derived as an implicature only after the literal meaning is processed and rejected as false. The rejection of the literal meaning, together with the assumption of the speaker's cooperative behaviour, would precisely trigger the inference-driven search for the speaker's intended meaning. Conversely, according to the direct access view (Gibbs, 1994), when seen in realistic social contexts, metaphorical meaning is directly accessed, with no need to pass through the interpretation (and rejection) of the literal meaning. Sharing the same assumption as the direct access view, Relevance Theory (Sperber & Wilson, 1995/86) sees metaphor - together with other non-literal expressions - as a case of loose use of language derived via the formation of ad hoc concepts (Carston, 2002; Wilson & Carston, 2007), occasion-specific concepts constructed online by narrowing and broadening the lexically encoded concepts. In this view, literal meaning plays no role but providing a clue for the interpretation process which is claimed to be inferential, relatively quick and automatic enough to readily allow the hearers to process metaphor in the explicature.

Intermediate models have also been proposed with respect to accessing literal meaning during metaphor interpretation. According to Giora's (1997, 2008) Graded Salience Hypothesis, metaphor interpretation is derived through parallel or sequential processing depending on the salience of the metaphorical expression: conventional metaphors are first processed both literally and metaphorically because the literal and metaphorical meanings are both salient; novel metaphor instead are interpreted through a sequential process in which, being more salient, the literal meaning is initially processed.

Interestingly, the theoretical discussion has recently developed and somewhat overcome the traditional approach to metaphor comprehension. On the one hand, metaphor interpretation is no more conceived dychotomically as either involving literal meaning or not: rather, the literal meaning of the metaphor is accessed in the early processing steps, and may sometimes linger (Carston, 2010; Carston & Wearing, 2011). Experimental evidence supports this view (Rubio-Fernández, 2007; Weiland, Bambini, Schumacher, 2014). On the other hand, some authors hypothesized that metaphor interpretation is not the same for all cases and for everybody, but it could be influenced by the type of metaphor, as well as personal, linguistic and socio-demographical factors. In particular, Carston (2010) and Carston & Wearing (2011) suggested that two different processing routes are possible, and that which of the two processing routes for metaphor understanding is taken depends on the interpretive demands. The first route is the default route, based on ad-hoc concept construction in the explicature. The second route assigns a greater role to literal meaning, which is maintained, meta-represented and subjected to a more reflective pragmatic inference. When the second route is undertaken, the activated literal meaning is attended to and used to derive a set of weak implicatures and other implications that lead to the interpretation of the metaphor. Crucially, the switch to the second route is claimed to occur under increased processing efforts. A typical case is the interpretation of extended metaphors, such as *'Depression, in Karla's experience, was a dull, inert thing - a toad that squatted wetly on your head until it finally gathered the energy to slither off'* (Heller, Z., *The Believers* quoted in Carston & Wearing, 2011, p. 295). In such expressions, the literal meaning of each chunk is said to prime the literal meaning of the subsequent chunk, overall resulting more accessible than the construction of several ad hoc concepts. However, the authors generalize this possibility also to instances of metaphors that are not particularly extended but still more opaque and less accessible, such as *'His ego is a flyblown globefish'* (Carston & Wearing, 2011, p. 309). In addition, they do not exclude the possibility that individual differences might occur with respect to which of the two routes is undertaken (p. 310) .

Following a compatible line of research, Gibbs (2010) and Gibbs & Colston (2012) have highlighted the need for theoretical accounts to include variation in the way a metaphor is processed and understood. They argue that metaphor interpretation is indeterminate: the experience of interpreting a metaphor is not universal, but it changes across individuals depending on several factors contributing to variations in the required cognitive effort. These factors concern linguistic and experimentally-related issues (e.g., metaphor familiarity/novelty and experimental methods used) as well as people's characteristics such as sex, culture, geographic origin and social status.

To summarize, metaphor interpretation has been traditionally theorized in terms of involvement or not of the sentence literal meaning. Recent theoretical positions, however, seem to push ahead this traditional debate and pave the way to the possibility that metaphorical meaning may be differently processed depending on the (variable) cognitive effort induced by the kind of metaphor used (e.g., more or less conventional, more or less extended) and at the same time by the comprehender's individual characteristics.

1.2 Metaphor Processing

ERP Correlates

Compared to the processing of literal language, metaphor processing requires more cognitive effort and this reflects in differential brain correlates. Neuroimaging studies have consistently shown that metaphorical expressions elicit increased brain metabolic activity (e.g., Bambini, Gentili, Ricciardi, Bertinetto, Pietrini, 2011; see Bohrn, Altmann, Jacobs, 2012 for a review). In electrophysiological research, two main Event-Related brain Potential (ERP) components are known to be involved in metaphor comprehension, namely the N400 and the P600/LPC (see Rataj, 2014 for a review).

The N400 is a negative-going waveform peaking at around 400ms after stimulus onset, with a typical centro-parietal distribution. This component has been first associated with semantic anomalies (Kutas & Hillyard, 1980; see Kutas & Federmeier, 2011 for a review). Today, the N400 amplitude is a reliable index of lexical semantic processing, sensitive to semantic incongruities and contextual expectations. As for metaphor processing, effects on the N400 component have been consistently reported: compared to literal language, metaphorical expressions elicit larger N400 amplitudes (e.g., Arzouan et al., 2007a; Arzouan, Goldstein, Faust, 2007b; Coulson & Van Petten, 2002, 2007; De Grawve, Swain, Holcomb, Ditman, Kuperberg, 2010; Lai et al., 2009; Pynte, Besson, Robichon, Poli, 1996; Weiland et al., 2014). In addition, the N400 amplitude has been found to vary as a function of the characteristics of the metaphor, notably its degree of conventionality and the presence/absence of context. Novel metaphors elicit larger N400 effects than conventional metaphors (e.g., Arzouan et al., 2007b; Coulson & Van Petten, 2002; Lai et al., 2009) and it has also been shown that a supportive context enhances metaphor interpretation reducing the N400 amplitude (Bambini, Bertini, Schaeken, Stella, Di Russo, 2016a). Overall, the N400 component elicited by metaphor interpretation has been functionally interpreted as indexing the extra-effort required during the lexical access phase of a metaphor, when the hearer retrieves the words using context-based expectations, and the search for the figurative meaning starts.

The P600/LPC component is a positive-going waveform that peaks at around 600ms (i.e., P600) after stimulus onset or later (i.e., LPC). Briefly, the P600 was initially associated with syntactic processing difficulties (e.g., Hagoort, Brown, Groothusen, 1993), and it has been then reported also for syntactically correct but semantically anomalous sentences (e.g., Kolk, Chwilla, van Hertnen, Oor, 2003) and in response to plausible but unexpected words (Federmeier, Wlotko, De Ochoa-Dewald, Kutas, 2007). More recently, P600/LPC effects have been reported for the interpretation of non-literal uses of language, such as metonymy (e.g., Schumacher, 2011), and – with a more frontal distribution - idioms (Canal, Pesciarelli, Vespignani, Molinaro, Cacciari, 2017) and irony (Spotorno, Cheylus, Van Der Henst, Noveck, 2013). For this reason, it has been argued that this late positive component might reflect not only syntactic reanalysis/sentence wrap-up but also pragmatic reanalysis (Regel, Meyer, Gunter, 2014). For metaphor, compared to the results on the N400 component, the findings on the P600/LPC are less consistent. Some studies have reported a positivity following the N400 for metaphorical than literal expressions (e.g., Bambini et al., 2016a; De Grauwe et al., 2010; Weiland et al., 2014). Some other studies have reported only a N400 effect for metaphor as compared to literal language (e.g., Arzouan et al., 2007a; Goldstein, Arzouan, Faust, 2012). This inconsistency has been recently explained in terms of differences in experimental designs and associated demands across studies (Rataj, 2014; Weiland et al., 2014). When found, the P600/LPC effect has been linked to the interpretation phase. In particular, Weiland et al. (2014) and Bambini et al. (2016a) proposed an electrophysiological account of metaphor processing according to which the biphasic N400-P600/LPC pattern reflects two different processing stages during metaphor comprehension. The N400 would index the cognitive efforts underlying a process of lexical/semantic access as shaped by context. Instead, the P600/LPC would reflect the costs associated with the most pragmatic aspects in deriving a metaphorical meaning (e.g., derivation of the speaker's intended meaning, implicatures, etc.).

Finally, a few ERP studies reported a sustained negativity following the N400 effect for novel (Arzouan et al., 2007a; Rutter, Kröger, Windmann, Hermann, Abraham, 2012) and literary metaphors (Bambini et al., 2018). Overall, this late negativity has been interpreted as an indicator of long-lasting efforts in deriving the non-literal interpretation, reflecting the increased difficulty in processing the metaphors at stake. For literary metaphors, the sustained negativity has been taken as evidence of the effort in elaborating the multitude of meanings which may have started within the N400 time window (Bambini et al., 2018). As for novel metaphors, slightly different, albeit compatible, interpretations have been proposed. According to Rutter et al. (2012), the sustained negativity reflects the extra effort necessary to integrate two unrelated concepts and, given the same centro-parietal distribution as the N400 effect, they also hypothesized such late negativity to be likely a prolonged N400. Arzouan et al. (2007a) suggested that the late negativity indicates further attempts to integrate meaning in a non-literal way. Interestingly, they also suggested that the absence of a P600/LPC effect might be due to the overlapping negativity which would attenuate the positive effect. Following this line of explanation, Rataj, Przekoracka-Krawczyk and Lubbe (2018) argued that this would be so because novel metaphors simultaneously require a prolonged process of information retrieval from long-term memory and extended mapping between semantically distant concepts.

Individual variation and Schooling

The study of how individual variation impacts language processing has interested research since a few decades only. However, there is evidence suggesting that cognitive performance and language proficiency are intertwined. For example, it has been shown that more versus less skilled readers exhibit a different cognitive profile (Perfetti, 2007) and have differently efficient suppression mechanisms (Gernsbacher, 1993). Consistently, recent studies using EEG methods have demonstrated that aspects related to the speaker's linguistic expertise greatly influence the brain response to linguistic stimuli, such as language proficiency (Pakulak & Neville, 2010), verbal fluency (Federmeier, Kutas, Schul, 2010; Lee & Federmeier, 2011) and punctuation habits (Steinhauer & Friederici, 2001). Despite what mentioned above, most research on metaphor understanding has been conducted with a focus on stimulus characteristics, while mostly ignoring participants' characteristics. However, a few studies found that subjects with higher IQ (Kazmerski, Blasko, Dessalegn, 2003), better verbal abilities (Blasko, 1999) and analogical reasoning abilities (Trick & Katz, 1989) exhibit a deeper comprehension of metaphors and seem to better capture the relational relevance between the metaphor target and vehicle (Trick & Katz, 1989). Moreover, a set of more recent studies suggests that metaphor processing prominently involves executive functions, especially working memory. In two eye-tracking experiments, higher working memory individuals (versus lower working memory subjects) were less likely to look back to metaphorical sentences (Olkoniemi, Ranta, Kaakinen, 2016), spent more time reading the critical region triggering the metaphor but were overall faster in reading the complete sentence (i.e., faster total reading times) (Columbus et al., 2015). In addition, higher working memory individuals were faster and more accurate than lower working memory subjects at understanding and producing metaphors (Chiappe & Chiappe, 2007) and at recognizing the non-literal meaning of a metaphor in a metaphor interference effect task (Pierce, MacLaren, Chiappe, 2010). Using brain imaging, Prat, Mason and Just (2012) reported evidence that working memory and vocabulary size influence the brain activation patterns for metaphor understanding. Interestingly, they found that lower working memory individuals showed increased activation in episodic-memory retrieval regions than higher working memory participants. Finally, it is worth noticing that these findings are corroborated by clinical evidence showing that individuals with executive control deficits (e.g., subjects with schizophrenia) exhibit difficulties interpreting metaphors (e.g., Bambini et al., 2016b; Mashal, Vishne, Laor, Titone, 2013).

Overall, what emerges from these studies is that, not only metaphor processing is influenced by individual cognitive performance, but also that working memory plays a primary role. This involvement has been explained as due to the need to activate and maintain both literal and figurative interpretations in mind (Olkoniemi et al., 2016), or maintain relevant semantic information and inhibit the irrelevant one (Prat et al., 2012).

Working memory and, more in general, individual cognitive performance might not be the only factors influencing metaphor processing. A few studies also documented effects of schooling in the appreciation of literary texts (Burkett & Goldman, 2016; Peskin, 2010) and effects of education level specifically for metaphor interpretation (Champagne-Lavau et al., 2012). Peskin (2010) tested students' ability to recognize and process texts in poetic and prose forms. Beyond age, participants differred in terms of literary education received: 4th graders (aged 9-to-10) had never received formal literary education, while grade 8 (aged 13-to-14) and grade 12 (aged 17-to-18) students had received 4 and 8 years of formal literary education, respectively. Participants were asked to read the literary passages and think

aloud while trying to make sense of the texts. Results revealed that only 12th graders exhibited a deeper attendance to the poetic texts while 4th graders were not even able to categorize a poem as a poem. Based on this, the author concluded that attendance to poetic texts (e.g., awareness of literary devices in poetry) systematically changes depending on the amount of formal literary education received. In a similar paradigm, Burkett & Goldman (2016) asked undergraduate students with no formal training in literature study to read literary texts and assessed their interpretation using a think-aloud task, which was followed by a post-reading interview. Their results revealed that literary novices were more likely to provide literary interpretations only when specifically prompted to do so, that is in the interview task than in the think-aloud task. The authors interpreted this as evidence that literary novices are likely to generate less sophisticated literary interpretations than experts.
To the best of my knowledge, Champagne-Lavau et al. (2012) is the only study directly assessing the role of education level on metaphor interpretation. In this study, a semantic judgement task was used to assess the interpretation of conventional and non-conventional metaphors (vs. literal sentences) in two groups of older participants with higher and lower level of formal education (i.e., respectively 12.5 and 6.7 years of schooling), who were also assessed for working memory capacity. Their results revealed a significant correlation with education level, but not with working memory: the higher participants' education level, the higher their accuracy rates in the metaphor task, thus arguing in favour of the idea that education level modulates metaphor interpretation.

Taken together, the studies above suggest that both education and individual cognitive performance are likely to model the cognitive demands imposed by metaphor interpretation. However, it is clear that a systematic line of research directly addressing this issue with more fine-grained methods and measures is still missing. The existing experimental data have provided some indications towards this direction, but we still ignore if and how individual variability in cognitive performance interacts with education and how this affects the electrophysiology of metaphor processing.

1.3 The present Study

The main aim of the present study is to investigate the role of education and a set of cognitive abilities in metaphor processing. In particular, the study wants to assess whether and how education and individual cognitive performance impact the electrophysiological correlates of novel metaphor interpretation. In order to do so, the EEG response to metaphor comprehension and literal controls was recorded from two groups of participants, academic and non-academic students. The two groups were matched for age and gender but differred in terms of education type: participants with an Academic Education (AE) were University students, whereas those with a Non-academic Education (N-AE) never attended University programmes and were attending vocational courses. To better characterize the two groups beyond the type of formal education and mere number of schooling years, measures of reading and writing habits were also taken. To the best of my knowledge, this is the first study directly testing the role of education type in metaphor interpretation and its ERP correlates as well as the role of cognitive abilities in the EEG response to metaphor.

In light of previous findings, a battery of neuropsychological tests was designed for the assessment of verbal working memory capacity, verbal IQ and semantic fluency. Including the neuropsychological battery in the study design allowed us (i) to investigate any correlation between the selected set of cognitive abilities and the ERP correlates for metaphor processing; and (ii) to assess and possibly disentangle the role of education type and individual cognitive variation in metaphor processing, i.e., to what extent individual differences affect processing beyond the effect of education.

As for the ERP correlates, the following is predicted: for the AE group, based on previous findings typically carried out on university students, it is expected to replicate the standard N400/P600-LPC biphasic pattern for metaphor interpretation, respectively linked to lexical access shaped by context and pragmatic interpretation (e.g., Weiland et al., 2014 and Bambini et al., 2016a). For the N-AE group, a different ERP pattern is expected, indicating higher efforts in metaphor interpretation. Differences may be due to increased difficulties at the lexical/semantic or inferential level, or alternatively to different processing strategies. There are only a few studies investigating the ERP response to linguistic stimuli as modulated by factors that might be closely related to education (e.g., language proficiency and punctuation habits, respectively in Pakulak & Neville, 2010 and Steinhauer & Friederici, 2001) and these all found consistent modifications of the ERP response depending on the subjects' characteristic at stake. However, since there are no studies on the role of education type in the ERPs response to metaphor, it is hard to develop more explicit experimental predictions.

With respect to individual cognitive performance as measured in this experiment, given previous behavioural findings, working memory is predicted to likely modulate the ERP response to metaphor. More in detail, previous studies found that higher working memory individuals exhibited facilitations in accessing the metaphorical meaning: in Columbus et al. (2015), they had faster total reading times and, in Olkoniemi et al. (2016), they were less likely to look back at the sentence. Overall, this suggests that working memory might enhance accessing the lexical meaning of the vehicle for selection of the metaphorically relevant properties. Within the EEG literature on metaphor understanding, this process has been interpreted as reflected by the N400 (Weiland et al., 2014; Bambini et al., 2016a). Higher working memory individuals may thus exhibit smaller N400 effects (lesser effort associated with lexical access) than lower working memory participants.

2. Methods

2.1 Participants

A total of 55 participants took part in the experiment (mean age = 25,01; SD = 3,96; age range: 19-35; 27 females). They were right-handed (scores > 85 according to the Edinburgh Handedness Inventory; Oldfield, 1971) native speakers of Italian, with normal or corrected-to-normal vision, no reading difficulties and no history of neurological/ psychiatric disorders. Participants were recruited according to their education type: 30 participants (mean age = 26,3; SD = 2,98; 14 females) were university students attending BA, MA or PhD programmes at the University of Pisa and Scuola Normale Superiore (mean years of Education = 19,6; SD = 2,76), whereas 25 (mean age = 23,48; SD = 4,47; 13 females; mean years of education = 12,64; SD = 1,60) completed the ten-years Italian mandatory schooling programme but never attended university courses and, at the moment of the experiment, were working in the constructions trade or were enrolled in a hairstylish/beautician vocational course. The group of 30 university students formed the Academic Education group (AE), while the 25 non-university students constituted the Non-Academic Education group (N-AE).

To better characterize the 2 experimental groups, participants were also administered a survey aimed at characterizing intellectual habits beyond mere years of schooling, assessing reading and writing habits, taken from Acheson, Wells & MacDonald (2008). For each section, participants were asked to approximately quantify - from 0 to 7 or more - the amount of hours per week they usually spent at - respectively - reading and writing some categories of documents. The categories included in the reading habits section were: *magazines, newspapers, e-mail, social networks, web-pages/sites* (excluding social networks), *novels* and *essays.* The categories included in the writing habits section were: *personal writing* (e.g., diary), *e-mail, social networks, creative writing* (e.g., poems, tales) and *material related to work activity* (e.g., notes, reports, etc.). The AE group spent significantly more hours per week than the N-AE group at both reading (AE: Mean = 33.59, SD = 14; N-AE: Mean = 19.52, SD = 6.59; W = 504.5; $p<0.0001$) and writing (AE: Mean = 18.03, SD = 8.10; N-AE: Mean = 12.21, SD = 5.98; W = 437.5; $p<0.05$), indicating that the two groups differed not just in terms of years of schooling but also in intellectual habits.

The present study was approved by the local Research Ethics Committee (Comitato Etico Area Vasta Nord Ovest, Azienda Ospedaliero-Universitaria Pisana). Written informed consent was obtained from every participant prior to the beginning of the experiment. Participants were paid for their time and none of them was aware of the goal of the experiment.

2.2 Stimuli

In the ERP paradigm, the experimental material was composed of a set of 80 pairs of not-lexicalized metaphors and literal counterparts (tot. 160) and a set of 160 filler items in Italian. The set of 80 pairs was composed of a subset of 42 pairs adapted from Bambini, Ghio, Moro & Schumacher (2013) and a subset of 38 newly created pairs (see Appendix B for a comprehensive list of the material used). All metaphors were nominal expressions of the form *X(s) is/are Y(s)* embedded in one-sentence context (e.g., *In hard times hopes are stars that light the soul up*[4]). The correspondent literal expressions were created by maintaining the vehicle and manipulating the topic of the metaphor and the context such

[4]Original item in Italian: *Nei momenti difficili le speranze sono stelle che illuminano l'anima*.

that a literal interpretation was obtained (e.g., *Those lights in the night sky are stars of distant galaxies*[5]). The 42 pairs from Bambini et al. (2013) were originally expressions of the form *Those Xs are Ys* and were adapted for the present experiment by enriching the sentential context to balance the number of words across sets (mean number of words = 11.17; SD = 1.53).

The material from Bambini et al. (2013) was already rated for meaningfulness, difficulty and familiarity (see Table 4). In order to ensure that the newly created materials were similar to the ones from Bambini et al. (2013), a questionnaire identical to the one used in Bambini et al. (2013) was used to rate stimuli for the same psycholinguistic variables used in Bambini et al. (2013). The questionnaire was administered online to 49 participants (mean age = 21,69; SD = 1,38). Participants were asked to rate on a 1-to-5 Likert scale the meaningfulness (i.e., how meaningful the sentence was) and difficulty (i.e., how difficult it was to judge the meaningfulness of the sentence) as well as the familiarity (i.e., how familiar the expression was to them/ frequency of experience) of the presented sentences. Mean values are reported in Table 4. Inferential non-parametric statistics showed that metaphorical expressions were as meaningful (W = 104.687; p = n.s.) and difficult (W = 80.044,5; p = n.s.) as literal expressions and, as expected, they were less familiar than the literal counterparts (W = 150.677; p< 0.001), in line with the set from Bambini et al. (2013).

Table 4. Mean(standard deviation) values for meaningfulness, difficulty and familiarity of the experimental material per item set (Bambini et al., 2013; new set; complete set) and condition (Literal, metaphorical).

	Materials from Bambini et al. (2013)		**New materials**		**Complete set**	
	Literal	Metaphorical	Literal	Metaphorical	Literal	Metaphorical
Meaningfulness	4.09(0.52)	3.31(0.59)	3.94(1.24)	3.95(1.05)	4.02(0.58)	3.16(0.62)
Difficulty	1.58(0.40)	1.97(0.42)	1.81(1)	1.81(0.89)	1.69(0.44)	1.89(0.40)
Familiarity	3.63(0.60)	2.69(0.98)	3.17(1.37)	2.98(1.25)	3.41(0.68)	2.83(0.83)

[5]Original item in Italian: *Quelle luci nel cielo notturno sono stelle di galassie lontane.*

2.3 Procedure

Participants were tested in one 2-hour experimental session, with the survey on reading and writing habits and the battery of neuropsychological tests administered respectively before and after the EEG recording. Participants were tested individually and were seated in a comfortable chair in the EEG Laboratory of Scuola Normale Superiore. The EEG session started with the presentation of written instructions followed by a short training phase aimed at familiarizing participants with the experimental procedure and task. The EEG recording lasted about 45/50 minutes. Materials were presented on a 19 inches computer screen situated at 100 cm in front of participants' eyes. Each trial started with a 750ms fixation cross displayed at the centre of the screen and followed by the experimental sentence presented word-by-word with a 600ms ISI and a 300ms blank screen to limit anticipatory ERP responses. ERPs were time-locked at target words (e.g., *In hard times hopes are **stars** that light the soul up*).

To ensure participants' attention, on 1/3 of the stimuli, at the end of the trial, participants had to perform a word matching task in which they had to indicate by button press which word out of two best fitted with the previous sentence, as previously done in other experiments on metaphor (e.g., Bambini et al. 2016a). The two lexical alternatives provided were displayed on the left and right of the screen for a maximum time window of 3000ms. If participants' responses exceeded the given time window, a new trial was presented. The two alternatives consisted of one related and one unrelated word: the related word always depicted a feature of the metaphor vehicle relevant for the interpretation of the metaphor, while the other word was always a totally unrelated alternative (e.g., *glare* vs. *switch* for the metaphorical item *In hard times hopes are stars that light the soul up*; see Appendix B for a complete list). The position in the screen (i.e., right/left) of the related and unrelated word was randomized across trials, as was the presentation of the trials. Both response times and accuracy were measured.

Participants were given two regularly scheduled breaks and were asked to refrain from blinking and moving their eyes during the word-by-word presentation of the sentences.

Participants were randomly assigned to one of two lists such that the same subject never read the same item in the two experimental conditions (i.e., metaphoric vs. literal). Each list was composed of 40 experimental pairs and 80 non-metaphorical filler items with the same syntactic structure as the experimental items (i.e., *X(s) is/are Y(s)*). Thus, the paradigm used in the EEG session was a 2x2 mixed Latin square design, with Education Type (*Academic vs. Non-Academic*) as the between-subjects variable and Condition (i.e., *metaphor vs. literal*) as the repeated, within-subjects, measure. Participants' ERP amplitude and the behavioural measures (i.e., accuracy and response times) represented the dependent variables.

2.4 Neuropsychological Measures

A computerized battery of cognitive/neuropsychological tests was administered to participants using Psychtoolbox 3 (Kleiner et al., 2007) toolbox for Matlab® (The MathWorks, Inc, Natick, US). The battery was administered after the EEG recording with an administration time ranging from 30 to 45 minutes overall. It was composed of the following 5 tests:

- *Verbal Working Memory Capacity* (WM), assessed through a sentence span task originally developed by Daneman and Carpenter (1980) and recently adapted in a larger Memory Battery (Lewandowsky, Oberauer, Yang, Ecker, 2010). Participants had to perform a dual task: they had to judge as true or false a statement appearing at the centre of the screen (e.g., *all women wear glasses*) and had to remember a consonant displayed after each statement for later recalling. Recalling occurred after a set of statements whose number varied from 3 to 8. The test gained a score between 0 and 1.
- *Verbal Fluency* (Fluency), measured with a test (Huff, Corkin, Growdon, 1986; Reverberi, Capitani, Laiacona, 2004) in which participants were given 60s time to list aloud all the items they could think of in relation to a given category. The categories used were: Fruit, Animals and Colours. This test was aimed at collecting a measure of participants' semantic network. Scores were obtained by counting the number of correct responses (1 score for every correct response).
- The *Test d'Intelligenza Breve* (TIB) (Sartori, Colombo, Vallar, Rusconi, Pinarello,1997) is an Italian adaptation of the National Adult Reading Test - NART (Nelson, 1982) and was used to obtain an estimate of Verbal IQ (VIQ). Participants were asked to read

aloud a set of 60 words, half of which had irregular accentuation in Italian. The final estimate of VIQ is obtained from the test score moderated by Age and Education.

- *Autism Quotient* (AQ) was tested using the Italian adaptation of Baron-Cohen, Wheelwright, Skinner, Martin, Clubley, 2001 (Ruta, D. Mazzone, Mazzone, Wheelwright, Baron-Cohen, 2012) to collect a measure related to theory of mind. Participants were asked to indicate how much they agreed on a four-point scale with a series of 50 statements concerning the 5 following sub-areas: Attention to details, Attention switch, Social Skills, Communication and Imagery. Higher scores are an indication of stronger autistic traits, with a score of 30 marking the clinical threshold.
- The *Author Recognition Test* (ART) (Acheson et al., 2008; Stanovich & West, 1989) was used to collect a formal index of exposure to reading. This was included in the battery because data on reading and writing habits (see the *Participants* sub-section) were based on a survey and, as such, did not provide a formal measure. On a list of 90 proper names (45 famous book writers), participants had to identify as many authors as they were sure to know. Here the English version (Acheson et al., 2008) was adapted by including the most famous contemporary Italian and international authors according to Italian bestseller charts since the last 10 years. Non-author names were taken from the editorial board of international Journals in neuroscience.

2.5 EEG recording and data pre-processing

The EEG was recorded from the scalp using 60 electrodes mounted in an appropriately sized cap (EasyCap Brain Products) according to the 10-20 International System: Fpz, Fp1/Fp2, Af3/Af4, Af7/Af8, Fz, F1/F2, F3/F4, F5/F6, F7/F8, Fc1/Fc2, Fc3/Fc4, Fc5/Fc6, Ft7/Ft8, Cz, C1/C2, C3/C4, C5/C6, T7/T8, Cpz, Cp1/Cp2, Cp3/Cp4, Cp5/Cp6, Tp7/Tp8, Tp9/Tp10, Pz, P1/P2, P3/P4, P5/P6, P7/P8, Poz, Po3/Po4, Po7/Po8, Oz, O1.

Horizontal eye movements were monitored using one electrode placed at the outer canthus of each eye. Vertical eye movements were monitored using two electrodes placed respectively over and beneath the right eye. Scalp electrode impedances were kept below 5 kΩ. Data from all scalp electrodes were referenced on-line to an electrode placed close to the vertex, and later off-line referenced to the mathematical average of the left and right mastoids. The raw EEG signal was collected in AC current with a low cut-off filter (time constant 10s) at a sampling rate of 500 Hz and was amplified using Brain Amp® amplifiers (Brain Products GmbH, Gilching, DE).

The EEG data were pre-processed using BrainVision Analyzer 2.0.3 (Brain Products GmbH). For each participant, epochs (from -600 to 1500ms) time-locked to the onset of the target word in the metaphorical and literal conditions were extracted. An independent component analysis (e.g., Groppe, Makeig, Kutas, 2008; Mennes, Wouters, Vanrumste, Lagae, Stiers, 2010) was carried out on the obtained epochs to identify and remove artifacts due to blinks and horizontal eye movements. The independent component analysis was conducted in a semi-automatic modality: the maximum amplitude range allowed in each epoch was fixed at 150 µV; the remaining epochs were visually inspected and corrected when contaminated by artifacts. After ICA correction, artifact inspection was conducted again to identify any residual artifact and the affected epochs were rejected. The average rejection rate was 6%. Participants with more than 40% rejection rate were excluded from final data analyses. Based on this criterion, 5 participants (3 AE; 2 N-AE) were excluded from statistical analyses. ERPs were computed for 1450ms after the onset of the target word relative to a -200ms stimulus baseline.

ERP waveforms were measured within specific time windows determined on the basis of previous literature on the ERPs correlates of metaphor interpretation and visual inspection of individual and group averages. Effects on the N400 and P600 components were expected, but, given the great variability in the literature, visual inspection was used to select the time windows of interest.

2.6 Statistical Analyses

For each selected ERP time window, statistical analyses were carried out on mean voltage amplitudes using ANOVAs with repeated measures and including two levels of Condition (i.e., metaphoric, literal), three levels of Longitude (i.e., Anterior, Central, Posterior) as well as the between-subjects factor, Education Type, with two levels (i.e., Academic, Non-Academic). Laterality was not included in these analyses for two main reasons: first, no lateralized effects were expected; second, visual inspection of the waveforms did not reveal any lateralization of the ERP components of interest. Within each time window, omnibus ANOVAs were performed on a subset of 30 electrodes (excluding peripheral electrodes) and were followed by additional ANOVAs conducted in a step-down fashion and/or, when needed, separate analyses per group to better characterize group differences or to break down significant between-groups interactions. Greenhouse-Geisser correction was applied to all ANOVAs with greater than one degree of freedom.

Between-groups analyses were also carried out on behavioural data from the ERP session using ANOVAs and including two levels of Condition (i.e., metaphoric, literal) and the two-level-between-groups factor (i.e., Academic and Non-Academic Education). Participants' accuracy and response times represented the dependent variables.

Two statistical analyses were carried out with the data collected through the battery of neuropsychological measures. First, for each of the cognitive measures included in the battery, inferential non-parametric statistics (i.e., Wilcoxon signed rank test) comparing the test scores between AE participants and N-AE participants was performed to estimate group differences. Second, the neuropsychological measures were used to investigate the correlation between Education type, cognitive abilities and the ERP components elicited by metaphor comprehension. To this purpose, a multiple regression analysis was carried out for each time window on the amplitude difference (i.e., metaphor minus literal) between the two conditions for each participant over a cluster of centro-parietal electrodes (C1, Cz, C2, Cp1, Cpz, Cp2, P1, Pz, P2) in the 300-450ms time window (i.e., N400 model) and a cluster of frontal electrodes (Af3, Fp1, Fpz, Fp2, Af4) in the 450-600ms time window (i.e., P600 model), respectively. In both models, Group (i.e., AE vs. N-AE) and participants' centred mean scores for each of the five tests (i.e., WM, Fluency, AQ, ART, VIQ) were used as predictors, as well as the resulting between-groups interactions. Given the exploratory nature of such analyses and the absence - to the best of my knowledge - of previous experimental evidence, a backward method based on AIC values was applied to determine the influence of each predictor within the model. Multicollinearity (VIF and mean VIF) and auto-correlation (Durbin-Watson test) diagnostics were carried out for each of the two models.

All statistical analyses were conducted using R software (R Core Team, 2015). All values from all statistical analyses, including the non-significant effects, are reported in tables VI and VII at the end of the chapter.

3. Results

3.1 Neuropsychological assessment: Results

Statistical analyses on participants' scores in each of the 5 tests revealed that the 2 groups significantly differed in all cognitive measures but Autism Quotient (AE: mean = 29.66; SD = 6.08; N-AE: mean = 28.40; SD = 3.94; W = 0.97, p = n.s.). AE participants had higher scores in the Sentence Span (SS) task used as a measure for WM (mean = 0.75; SD = 0.13) than N-AE subjects (mean = 0.52; SD = 0.22) and the difference between groups was highly significant (W = 601; $p < 0.0001$). A similar pattern emerged with performances on the Verbal Fluency task, with AE participants achieving significantly higher scores than N-AE subjects (AE: mean = 50.23; SD = 9.98; N-AE: mean = 43.92; SD = 9.03; W = 504.5; $p < 0.05$).

The AE group significantly differed from the N-AE group in the task used to collect a measure for Verbal IQ, with higher scores for AE than N-AE participants (AE: mean = 114.33; SD = 13.55; N-AE: mean = 98.73; SD = 5.83; W = 740; $p < 0.0001$).

Finally, as expected, AE participants obtained higher scores than N-AE subjects in the ART test (AE: mean = 21.63; SD = 9.43; N-AE: mean = 8.6; SD = 6.11) and inferential statistics confirmed that the difference between groups was highly significant (W = 663; $p < 0.0001$).

3.2 Behavioural Results

For the AE group, the mean frequency of correct selection of the related-word option was 78.29% and 70.04% in the metaphoric and literal condition respectively. N-AE participants selected the correct option 79.76% and 68.29% of the time in the metaphorical and literal condition respectively. The between-groups ANOVA revealed that, while there were no group-related effects (Group: $F(1,47)<1$, ns; GroupXCondition: $F(1,47)<1$, ns), all participants were significantly more accurate in the metaphorical than the literal condition ($F(1,47)=22.50$; $p<0.005$).

Mean response times (ms) of the AE group were 1252(382.22) and 1518(480.9) for the metaphoric and literal conditions respectively, while N-AE's mean response times (ms) were 1453(491.79) in the metaphoric condition and 1432(478.52) in the literal condition. The between-groups ANOVA revealed no significant differences between conditions ($F(1,47)<1$, ns) and Groups ($F(1,47)<1$, ns) nor any significant interaction ConditionXGroup ($F(1,47)=1.15$, ns).

Overall, these results suggest that participants performed the task correctly, attended the presented materials while their EEG was recorded and, independently of their education type, comprehended the experimental stimuli.

3.3 ERP Results

The ERP data to the target word in metaphorical and literal expressions and the relative scalp maps for the distribution of the effects for AE and N-AE groups in the time windows of interest are shown in Figures 7 and 8. Visual inspection of the waveforms and scalp maps revealed the following differences between groups. In the AE group, compared to the interpretation of literal expressions, metaphor interpretation elicited more negative ERPs with a centro-parietal distribution, followed by more positive ERPs over frontal sites. In the N-AE group, compared to the interpretation of literal language, metaphorical expressions elicited more negative ERPs that were more widely distributed over scalp sites. Such negative effect was not followed by any positive effect; rather the negativity persisted over centro-parietal electrodes (see Figure 8). On the basis of this inspection, the whole analysis was focused on the time window ranging from 300 to 600 ms and, to avoid issues related to the temporally overlapping of the components, two portions of the time window were tested independently: the early portion (300 to 450 ms) was used to investigate the N400 component, and the late portion (450 to 600 ms) to investigate the positive effect and the prolonged negativity.

300-to-450ms Time window

A first omnibus between-groups ANOVA (see table VI) over the subset of 30 electrodes in the 300-450ms time window revealed not only a significant effect of Condition ($F(1,48)=22.56$; $p<0.0001$) and Longitude ($F(2,96)=11.43$, $p<0.005$, $\varepsilon=0.70$), but also a significant GroupXLongitude interaction ($F(2,96)=8.14$, $p<0.005$, $\varepsilon=0.70$) and, more importantly, a significant GroupXConditionXLongitude interaction ($F(2,96)=3.57$, $p=0.05$, $\varepsilon=0.65$), suggesting that the topography of the N400 effect elicited by metaphor

interpretation differed between groups. Subsequent 2-Way ANOVAs conducted separately per longitude level confirmed the above pattern. There was a significant effect of Condition (−0.53 μV, $F(1,48)=6.05$, $p<0.05$) and a significant GroupXCondition interaction (AE: +0.04 μV, N-AE: −1.20 μV, $F(1,48)=6.92$, $p<0.05$) over anterior electrodes, with a significant negative effect in the N-AE group only (see Figure 7). Conversely, independently of education type, the amplitude values relative to the metaphorical condition significantly differed from those in the literal condition over both central (Condition: −1.04 μV, $F(1,48)=21.88$; $p<0.0001$; GroupXCondition: $F(1,48)<1$, ns) and posterior (Condition: −0.83 μV, $F(1,48)=17.58$; $p<0.0001$; GroupXCondition: $F(1,48)<1$, ns) electrodes.
To summarize, these results indicate that metaphor interpretation elicited a N400 effect and that Education type modelled its distribution over the scalp: centro-parietal in AE participants and more widely distributed in the N-AE group.

450-to-600ms Time window

The omnibus ANOVA over the subset of 30 electrodes in the 450-600ms time window (see table VI) revealed a significant effect of Condition ($F(1,48)=4.91$; $p<0.05$) and a marginal interaction GroupXCondition ($F(1,48)=2.99$; $p=0.09$). Since the above interaction approached significance and visual inspection of the waveforms suggested a frontal positivity in the AE group and a prolonged centro-parietal negativity in the N-AE group, additional ANOVAs were carried out separately per group to better identify and characterize this group-difference.

In the AE group, a 2-Way ANOVA conducted over the subset of 30 electrodes revealed a significant effect of Longitude ($F(2,52)=8.93$; $\varepsilon=0.59$; $p<0.005$) and a significant interaction ConditionXLongitude ($F(2,52)=8.02$; $\varepsilon=0.71$; $p<0.005$). A subsequent 1-Way ANOVA conducted on a cluster of frontal electrodes (i.e., Af3, Af4, Fp1, Fp2, Fpz) confirmed that such positivity effect elicited by metaphorical expressions as compared to literal ones was localized over frontal sites (Condition: +0.79 μV, $F(1,26)=4.85$; $p<0.05$).

Statistics on the ERP data from the N-AE group revealed a remarkably different pattern of results in the 450-600ms time window. The omnibus 2-Way ANOVA over 30 electrodes showed a significant main effect of Condition ($F(1,22)=7.92$; $p<0.05$).

Furthermore, since visual inspection of the waveforms suggested the absence of a positivity effect and the presence of a negativity one (see Figure 8), two 1-Way ANOVAs were conducted over (i) a cluster of frontal electrodes (i.e., Af3, Af4, Fp1, Fp2, Fpz) to exclude the positivity effect; and (ii) a cluster of centro-parietal electrodes (i.e., C1, Cz, C2, Cp1, Cpz, Cp2, P1, Pz, P2) to verify the negativity effect. The first 1-Way ANOVA over the frontal cluster confirmed no P600 effect in the N-AE group (Condition: $F(1,22)<1$, ns). Conversely, the second 1-Way ANOVA on the centro-parietal cluster confirmed that the negativity effect was statistically significant (Condition: −1.31 µV, $F(1,22)=10.07$; $p<0.005$).

Overall, these results suggest that, compared to literal expressions, the interpretation of metaphorical expressions elicits different electrophysiological patterns in the 450-600ms time window depending on Education type: while AE participants exhibited a frontal P600 effect, this was not the case for N-AE participants, who instead showed a prolonged negativity.

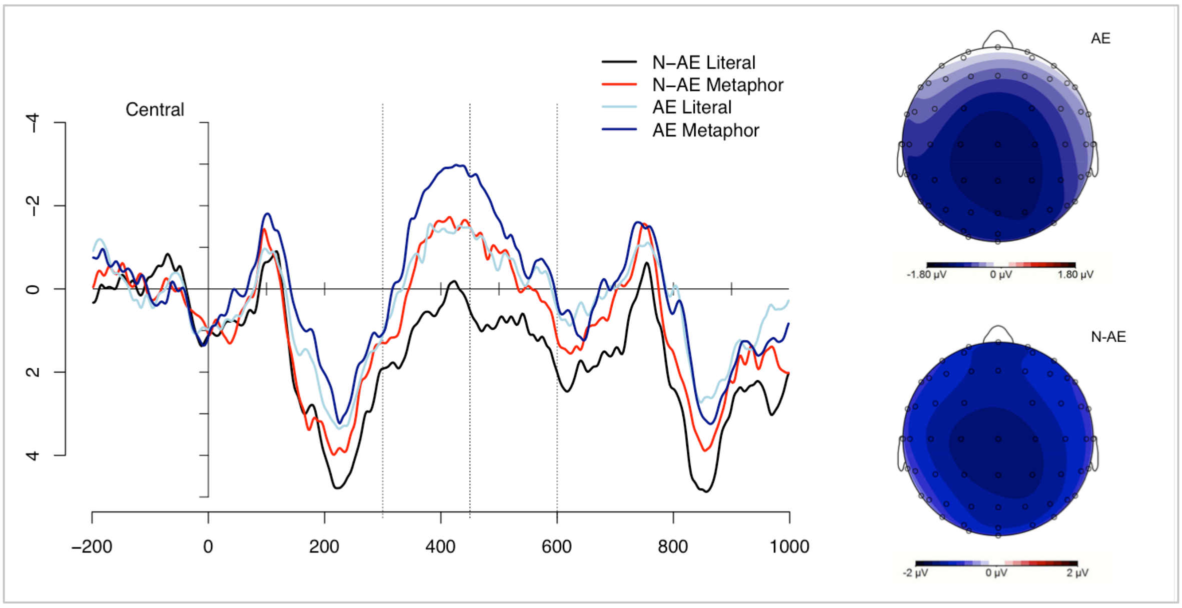


Figure 7: ERP waveforms for the metaphorical and literal conditions in the AE and N-AE groups at central electrodes. Dotted grey bars indicate the time windows of interest. Right panel: scalp maps of the mean average difference between metaphorical and literal conditions in the AE (top) and N-AE (bottom) group in the 300-to-450ms time interval.

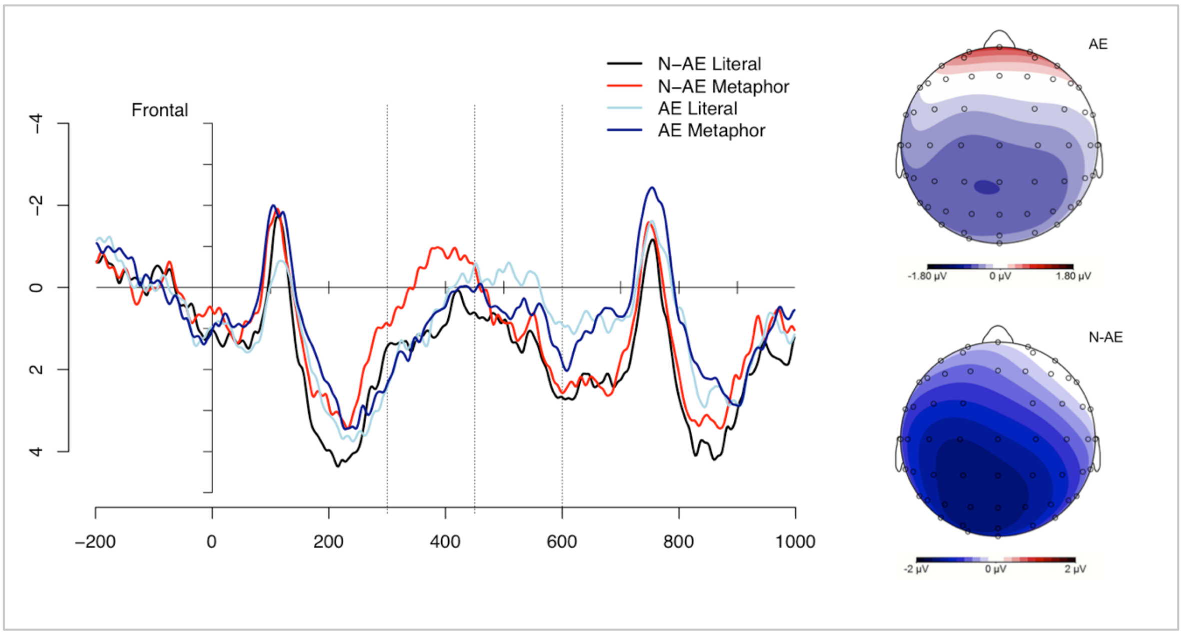


Figure 8: ERP waveforms for the metaphorical and literal conditions in the AE and N-AE groups at frontal electrodes. Dotted grey bars indicate the time windows of interest. Right panel: scalp maps of the mean average difference between metaphorical and literal conditions in the AE (top) and N-AE (bottom) group in the 450-to-600ms time interval.

3.4 Multiple Regression Analyses: Results

The resulting statistics of the final backward multiple regression N400 model are reported in Table VII, together with details on the diagnostic tests for multicollinearity and autocorrelation.

Overall, results revealed that Working Memory (i.e., SS), Exposure to reading (i.e., ART) and Education TYPE (i.e., Academic vs. Non-Academic) predicted the neural response to metaphor interpretation in the 300-450ms time window. Specifically, exposure to reading had the greatest impact on the modulation of the N400 average difference amplitude (β_i= .67; $p<0.001$): the higher the scores in the ART test, the smaller the N400 effect (Figure 9a). This was followed by Education type (β_i= −.50; $p<0.05$), for which a negative correlation emerged thus indicating that the higher the education type (i.e., AE Group) the more negative the N400 (i.e., the larger the N400). Finally, Working Memory - as indexed by the SS test - was found to be the last significant predictor (β_i= .39; $p<0.05$): similarly to the effect of Exposure to reading, the higher the performance in the WM test, the more positive the N400 averaged difference amplitude, thus the smaller the N400 effect (Figure 9b).

As for the multiple regression analysis conducted with the amplitude values in the 450-600ms time window, the statistics for the overall model revealed that the model itself was not significant ($\Delta R^2 = 0.10$; $F(5,44)=2.12$; p=n.s.) and the diagnostic tests revealed that the model was affected by multicollinearity (mean VIF = 2.75) and autocorrelation (D-W= 2.71; $p<0.05$).

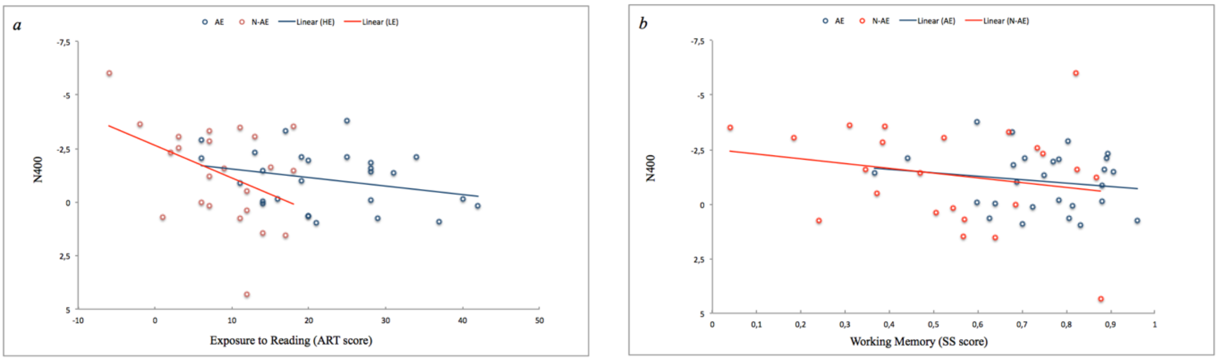


Figure 9: Correlation between the average difference amplitude (metaphorical – literal, in µV) over centro-parietal sites (C1, Cz, C2, Cp1, Cpz, Cp2, P1, Pz, P2) and Exposure to Reading (a) and Working Memory (b) in the 300-to-450ms time window.

4. Discussion

This study aimed at investigating the role of Education type and individual cognitive abilities in metaphor processing. To do so, the EEG response to metaphorical and literal expressions was recorded in individuals with an academic and a non-academic education and some measures of their cognitive performance were collected, by administering a neuropsychological battery that assessed verbal working memory, semantic fluency, verbal IQ, autism quotient and exposure to reading.

Statistical analyses revealed two main results. First, academic and non-academic students exhibited different ERP correlates to metaphor processing: while the AE group showed a biphasic N400/P600 pattern, with respectively a centro-parietal and a frontal distribution, a monophasic N400 pattern was found in the N-AE group, more largely distributed and prolonged over time. Second, the N400 amplitude for metaphor interpretation significantly correlated with participants' education type, but also with individual working memory capacity and exposure to reading. In addition to this, as expected, AE participants exhibited a more efficient cognitive performance than N-AE individuals given that they scored significantly higher in verbal IQ, exposure to reading, verbal working memory and verbal fluency tests. In what follows, first group differences are discussed as well as what might be the different underlying cognitive processes, considering several hypotheses (4.1), then the role of individual variation in cognitive performance (4.2).

4.1 Group differences between Academic and Non-Academic learners

Start from the first finding, different ERP response to metaphor understanding emerged depending on education type. As expected, the biphasic N400/P600 pattern exhibited by the AE group replicates previous ERP literature on metaphor understanding (e.g., Bambini et al., 2016a; De Grawve et al., 2010; Weiland et al., 2014; see Rataj, 2014 for a review). Following Weiland et al. (2014) and Bambini et al. (2016a), this biphasic response indexes the costs associated to, respectively, lexical access as shaped by context and further pragmatic interpretation. Based on this account, then, the N400 effect suggests that, compared to literal expressions, in metaphor comprehension, higher cognitive effort is required to access the lexical meaning and the relevant properties of the metaphoric vehicle depending on the constraints posed by the context. The P600 effect instead may reflect the effort spent for further processes, presumably related to pragmatically enrich the vehicle such that it can be meaningfully applied to the target of the metaphor and finally derive the speaker's intended meaning.

As for the distribution of this biphasic pattern, the N400 effect exhibited the typical centro-parietal distribution thus replicating previous findings (see Rataj et al., 2014 for an overview). The P600 component showed a frontal distribution. To this purpose, some points are noteworthy. First, as stated in the introduction, the findings of the P600/LPC for metaphor comprehension are much less consolidated than those relative to the N400 component. Second, the topography itself of the positive effect is still a matter of debate over the literature: it exhibited a centro-posterior distribution in Weiland et al. (2014) and a parietal distribution in Bambini et al. (2016a). Furthermore, De Grawve et al. (2010) reported an early LPC effect over central sites (exp. 1) and a later one over fronto-central (exp. 1) and posterior (exp. 2) sites. Overall, then, further experimental work would be needed to better clarify the topographical issues of the pragmatic P600/LPC for metaphor understanding. However, it is also worth noting that the finding emerged in this study is consistent with Canal et al. (2017) and Spotorno et al. (2013) who found a more frontally distributed positivity for other pragmatic phenomena, that is idioms and irony understanding respectively.

The N-AE group exhibited a different response: no positive effect was found in the P600 time window, but rather the N400 was prolonged in time. Differences between the two groups also concerned the scalp-distribution of the N400 effect, which was more broadly distributed in the N-AE group. Overall, these group differences confirm the very general prediction that education type modulates the ERP correlates for metaphor understanding. In particular, the centro-parietal negativity was elicited in a time window that followed the N400. As stated above, the N400 has been functionally interpreted as reflecting the costs associated with lexical access as shaped by the previous context, that is the context-guided search in semantic space for relevant properties of the vehicle which may be attributed to the topic (Bambini et al., 2016a; Weiland et al., 2014). Immediately following the N400, some previous studies reported late negative effects during the comprehension of novel (Arzouan et al., 2007a; Rutter et al., 2012) and literary metaphors (Bambini et al., 2018), and they were taken as reflecting extra and long-lasting effort in deriving the figurative interpretation. It was also noted that the late negativity would actually mirror prolonged efforts already starting in the N400 time window and extending to the P600 time window (see also Rataj et al., 2018). In line with such interpretations, the prolonged N400 pattern observed here can be interpreted as evidence for higher lexical/semantic processing efforts in the N-AE group than the AE one.

Interestingly, the behavioural results did not reveal any group-related differences, neither in accuracy nor in RTs. The only significant result was a main effect of condition with higher accuracy rates for metaphors than literal controls, which is likely due to the fact that the related word of the word matching task (e.g., *glare*) was more salient to the associated metaphorical (e.g., *[…] hopes are stars […])* than literal expression (e.g., *[…] those lights are stars […]).* Besides, this pattern is consistent with Noveck, Bianco and Castri (2001), who found adults' more correct responses to metaphors than literal controls. Thus, behavioural results suggest that both groups understood metaphors and argue against the idea that non-academic education impedes metaphor understanding. Rather, these data indicate that education type impacts metaphor interpretation on the underlying cognitive processing indexed in the ERPs correlates, and specifically point to a greater effort in the lexical/semantic phase. What exactly such greater efforts are associated with?

Hypothesis 1: Greater effort in context-driven lexical operations for N-AE

One possible explanation is that the prolonged N400 indexes N-AE individuals' **increased difficulties** during the manipulation of lexical elements in the context of the expression, which ultimately leads to metaphor interpretation. That is, compared to AE subjects, lexical access as shaped by context - which allows for further pragmatic operations of concept adjustment - might require higher and more prolonged cognitive efforts that result in a sustained N400 carrying over up to 600 ms. In discussing their ERP findings, Weiland et al. (2014) and Bambini et al. (2016a) propose an explanation that is consistent with a relevance-oriented theory (Wilson & Carston, 2007). So, for example, take a metaphor such as 'Hope is a star'. In order to understand this metaphor, it is first and foremost necessary to access the lexical meaning of the vehicle and integrate it with the context of the expression to figure out the link between 'hope' and 'star' and bridge the gap between these concepts. According to the ad hoc concepts account, some properties of the lexical concept STAR are relevant for the whole interpretation, for example BRIGHT, which would be further adjusted (i.e., narrowed and broadened) to form an ad hoc concept BRIGHT* that includes in its denotation not only stars but other 'enlightening' entities, such as hope. The final outcome would be an inferentially derived interpretation of the metaphor roughly paraphrasable as 'Hope enlightens people's life, motivates human beings in achieving their aims and overcoming difficulties, etc.'. Importantly, when first hearing 'Hope is a star', the lexical meaning of the vehicle is accessed and a process of contextual integration is initiated. This enhances accessing those properties of the lexical concept STAR which are relevant for the interpretation of the metaphor and which, according to the ad hoc concept account, will undergo pragmatic adjustment. Following Weiland et al. and Bambini et al. interpretation, the N400 would reflect the costs required by the first processing phase, in which readers access the lexical meaning of the vehicle and, based on context expectations, try to integrate it with the context. These results are compatible with the idea that AE subjects overcome the lexical access phase by readily integrating the context (i.e., the canonical N400 effect) such that relevant information of the vehicle concept is accessed and this allows them to move forward in the interpretive process to eventually adjust the metaphoric meaning of the vehicle (i.e., P600 effect). Conversely, N-AE individuals might experience greater difficulties at this context-driven lexical operation phase, thus triggering more cognitive effort at the lexical/semantic level, reflected in the long-lasting N400. In sum, the different ERP patterns between groups indicate higher processing effort for N-AE individuals that might stem from difficulties at the context-driven lexical access phase, which allows for relevant information of the vehicle to be accessed. Following this interpretation, these data suggest that education type models the processing costs for metaphor understanding by facilitating the lexical access of the vehicle as shaped by context.

There are some points worth noting with respect to what said above. First, interpreting the prolonged N400 in terms of greater difficulties in metaphor understanding at the lexical level may be consistent with the results on novel and literary metaphors, where long-lasting negativities were reported by Arzouan et al. (2007a), Rutter et al. (2012) and Bambini et al. (2018). Novel and literary metaphors differ in many respects but, compared to conventional metaphors, they both are cases in which processing is taxed because the relation between target and vehicle is much less accessible. It might be the case that the prolonged N400 effect observed here may last for several milliseconds beyond the time window under scrutiny, prolonging in the P600 time window (see Rataj et al., 2018).

Finally, interpreting the group differences in terms of N-AE greater processing efforts fits smoothly with the recent theoretical discussion developed by Gibbs (2010) and Gibbs & Colston (2012). They argued that metaphor accessibility might not be the same for everyone, but the cognitive effort underlying metaphor interpretation is variable and might also depend on people’s characteristics. These findings suggest that education type is one such factor and provide novel experimental evidence to this purpose. In addition, an interpretation of these ERP data along the lines of greater processing efforts for N-AE individuals stands out harmoniously with previous behavioural findings showing that the higher participants' education level the higher their accuracy in a metaphor task (Champagne-Lavau et al., 2012), and that the level of education attained contributes to a better appreciation of literary texts (Burkett & Goldman, 2016; Perskin, 2010).

Hypothesis 2: Different processing strategies for AE and N-AE

An alternative explanation of the findings revealed by the present study could be that the different ERP patterns in AE and N-AE subjects reflect **two different processing strategies**, along the lines of Carston & Wearing (2011) (see also Carston, 2010). As mentioned in the introduction of this chapter, these authors have proposed two processing routes for metaphor interpretation, namely (i) ad hoc concept formation (i.e., the standard, relatively costly, strategy) and (ii) a process in which the literal meaning lingers. The second route is one in which the literal meaning is meta-represented, maintained and used as material for a more reflective pragmatic inference such that metaphor understanding would be achieved as the result of several weak implicatures. The authors consider this process slower than ad hoc concepts formation, but also more convenient in terms of processing efforts. In fact, the system would switch to the second route when the online construction of ad hoc concepts taxes the individual's cognitive system beyond a threshold after which processing literal meaning - and deriving metaphor as part of the implicated meaning - is, ultimately, less demanding. In this proposal, extended metaphors are seen as the paradigmatic case for the second route. However, the authors speculated that the switch from the classic to the second route might occur 'whenever the local processing load reaches a level at which the effort of accessing or constructing an ad hoc concept is too great' (Carston & Wearing, 2011, p. 309), which might also depend on individual differences. The degree of accessibility of a metaphor might indeed be bound to people's cognitive resources too, which are likely to influence the processing load of metaphor interpretation. In addition, what these ERP findings first and foremost suggest is that there are differences in terms of ERP response to metaphor depending on education type. Thus, it might be the case that education type contributes determining the degree of accessibility of a metaphor such that one and the same metaphor such as '*Hope is star*' could be accessible enough for AE participants to form ad hoc concepts constructed on the fly, while it could require N-AE individuals to attend literal meaning and this would trigger the second processing route. If this was the case, then, the different ERP patterns in the two groups might be taken as indicators of the cognitive efforts associated with different interpretive modalities: the biphasic N400/P600 pattern in the AE group might index a processing mechanism along the lines of ad hoc concepts formation, and the prolonged negativity pattern in the N-AE group might reflect the costs associated to the second route. In this case, the prolonged N400 could reflect a process in which aspects of the literal meaning linger and are attended to. Consequently, following Carston & Wearing (2011) proposal, this would lead to an array of weak implicatures through which metaphorical interpretation is finally derived.

Notice that caution is needed here. The present experiment did not directly address any issue related to the role of literal meaning in metaphor interpretation and the design did not include any manipulation to this purpose. However, there is evidence in the literature that might support such an interpretation. Previous research suggests that literal meaning, or at least some aspects of it, seems to be active during the earlier phase of metaphor interpretation (Rubio-Fernández, 2007; Weiland et al., 2014 - exp. 1), and this raises the issue of when irrelevant aspects of the literal meaning are suppressed during metaphor interpretation. In addition, there is evidence for individual variation in the ability of suppressing irrelevant meanings (Gernsbacher, 1993) and in the ability of focusing on the false literal meaning of a metaphoric expression (Kazmerski et al., 2003). Gernsbacher (1993) showed that less skilled readers, as compared to more skilled readers, are less efficient in suppressing inappropriate word meanings: contrary to more skilled readers, for less skilled readers contextually inappropriate meanings (e.g., *ace* for *He dug with the spade*) were as activated 1 second later the presentation of the target item as they were immediately (i.e., 100ms). Interestingly, in this study, participants were U.S. Air Force recruits with a high-school diploma (only 20% of the sample had completed some college courses; p. 295). Thus, although this was not a study on the effect of education on suppression mechanisms, the present results may be due to N-AE participants paying more attention to the literal meaning of the vehicle. Furthermore, Kazmerski et al. (2003) showed that low IQ participants (versus high IQ ones) exhibit less interference effect - i.e., no response times difference and small N400 effect - in a metaphor interference effect (MIE) task. In a typical MIE task, people take longer judging a metaphor as literally false than scrambled sentences and this is traditionally taken as evidence for the automaticity of metaphor processing, without the need to go through the literal meaning first (see Glucksberg, 2003 for a review). The authors interpreted their findings as suggesting that high IQ subjects, but not low IQ, were facilitated in the early activation of metaphorically-relevant features. Consistently with this, in the present study, N-AE participants scored significantly lower in the IQ test than participants in the AE group. Overall, then, the studies above raise the possibility that different interpretive mechanisms, differently relying on literal meaning, occurred for the two groups of AE and N-AE participants. In the lexical access phase as shaped by context, AE individuals might have readily suppressed the metaphor irrelevant properties of the literal meaning of the vehicle (e.g., CELESTIAL BODY for STAR), activated/selected relevant information (e.g., ENLIGHT) and this might have enhanced metaphor interpretation, for example facilitating the formation of ad hoc concepts (e.g., STAR*). Conversely, N-AE individuals might have had difficulties suppressing the irrelevant properties or selecting the relevant ones and, as a result, remained anchored to the literal meaning of the metaphorical expressions or, at least, they did so longer than AE participants. Consequently, it is possible that their interpretation followed a path in which literal meaning was actually attended to and exploited for the understanding of a metaphor, along the lines of Carston & Wearing's (2011). In this light, hence, the sustained N400 effect might reflect N-AE individuals' greater attendance to the literal meaning of the metaphor. However, the present study does not provide direct evidence on whether the two groups of participants differently used literal meaning depending on their available processing resources. As such, although not implausible, the considerations above remain intriguing yet speculative to some extent. Further experimental work would be needed to verify this possibility.

4.2 Individual cognitive differences in metaphor comprehension

To what extent does the ERP variation discussed above depend on **individual variation** in cognitive performance, education type and/or the interaction between the two? This was the second research question addressed by this study. Results indicated that the N400 amplitude significantly correlated with education type, but also with participants' verbal working memory and exposure to reading, independently of education type. Importantly, the multiple regression analysis indicated that the group effect accounted for 50% of the variance in the N400 amplitude. This further confirms the ERP data as emerged in the two groups and suggests that education type plays a prominent role in modulating the ERP correlates of metaphor understanding.

Nonetheless, independent of group, both verbal working memory and exposure to reading positively correlated with the N400 amplitude. The positive correlation between the ART score and the N400 amplitude suggests that the higher the ART score, the more positive the N400 response to metaphor understanding. Hence, individuals with higher ART scores also exhibited a less pronounced N400 effect. This indicates that exposure to reading enhances metaphor processing by reducing the costs associated to lexical access as shaped by context. In other words, the more an individual is exposed to reading, the smaller the cognitive costs associated with the lexical semantic processing of the metaphoric vehicle. This result is not surprising because the access to lexical information during metaphor comprehension has to do with the ability to select the relevant meaning depending on context and, presumably, the semantic network of individuals who are more used to reading is more efficient compared to less skilled readers. A well-read semantic network may facilitate the activation (and subsequent selection) of the relevant meaning (see also Kazmerski et al., 2003 for a similar discussion). Moreover, this result is in line with Prat et al. (2012), who found that vocabulary size (a measure that is connected to our ART test) correlates with the brain patterns activated during metaphor interpretation. Finally, the present results suggest that exposure to reading predicts the N400 amplitude for metaphor and this is so independently of education type, even though individuals with academic education showed significantly higher ART scores than individuals with a non-academic education.

Consistently with the predictions of this study, statistical analyses also revealed that verbal working memory modulates the N400 amplitude elicited by metaphor interpretation. In particular, the higher participants' working memory scores in the SS task, the more positive the amplitude of the N400 effect. Hence, individuals with better working memory capacity also exhibited less pronounced N400. This suggests that, similarly to exposure to reading, working memory enhances lexical access shaped by context as needed to understand a metaphor. That is, the better the individual's working memory ability, the less the cognitive costs required to select the metaphor relevant property(ies), given the context. This fully fits with studies on metaphor interpretation where significant correlations between metaphor processing and working memory were found (e.g., Olkoniemi et al., 2016; Prat et al., 2012) and further corroborates the functional interpretation of the N400 as an index of those preliminary lexical operations needed to derive metaphorical meanings. It is in fact reasonable that the multiple properties of the metaphor lexical concepts are activated in working memory and that higher working memory capacity facilitates the activation and selection of the relevant metaphoric properties (see Prat et al., 2012 for a similar explanation). Again, the multiple regression results do not allow us to conclude that this is strictly dependent on education type, but statistics on the neuropsychological tests revealed that AE students exhibited significantly better working memory performance than N-AE ones.

In sum, the main result from the multiple regression analysis conducted for the present study is that education plays a massive role in modulating the ERP response to metaphor interpretation related to lexical operation mechanisms. However, variation in the ERP response is also explained by individual differences: both exposure to reading and working memory contribute to metaphor processing by reducing the cognitive costs associated to the lexical access phase as indexed by the N400 effect. Exposure to reading is an indicator of individuals' semantic networks. In metaphor interpretation, the relevant properties are selected from a range of properties of the lexical encoded concept which, in turn, are all activated during early processing. It is hence possible that the greater an individual's semantic networks, the more readily available those properties are in working memory.

2.5 Conclusion

Except for a few recent theoretical proposals, the traditional debate on metaphor understanding has always focused on how a metaphor is understood based on whether or not the processing of literal meaning is passed through. Experimental research has shown that, whatever the interpretive strategy, metaphor understanding is not for free. It does come at a cost, which varies because of figurativity per se (i.e., metaphor versus literal language) and depending on what the metaphor is (i.e., novel, poetic, etc.). However, the study of what metaphor processing involves because of its intrinsic features only informs us about one side of the coin and metaphor understanding might be even more intricated than this. Human beings differ in a variety of ways which include both endogenous (e.g., individual variation in cognitive performance) and exogenous factors (e.g., geographic origin, culture, life experience, education, etc.) as well as the interaction between the two. By focusing on the other side of the coin, the present study further opened up the whole picture and suggested that the cognitive costs underlying metaphor interpretation also vary depending on who the processor is and his/her education type. Unlike non-academic education, academic education seems to reduce the cognitive efforts required during the lexical access phase of metaphor processing, presumably enhancing the integration in context of the vehicle lexical meaning, which allows for relevant information of the vehicle to be accessed and pragmatically adjusted for further derivation of the speaker's intended meaning.

Table VI - Analyses of Variance for the N400 time window (300-450ms) and the P600 time window (450-600ms).

350-to-450ms time window					
	df	F	p	ε	Size effect (generalized eta-squared)
Condition	1, 48	22.56	< 0.0001***		0.036
Group	1, 48	0.46	0.60		0.006
Longitude	2, 96	11.43	0.0003**	0.70	0.039
GroupXLongitude	2, 96	8.14	0.002**	0.70	0.028
ConditionXGroup	1, 48	2.38	0.10		0.004
ConditionXLongitude	2, 96	1.99	0.15	0.65	0.002
ConditionXGroupXLongitude	2, 96	3.57	0.05*	0.65	0.004
Comparisons					
Anterior electrodes					
Condition	1, 48	6.05	0.01*		0.016
Group		1.15	0.28		0.020
ConditionXGroup	1, 48	6.92	0.01*		0.018
Central electrodes					
Condition	1, 48	21.88	<0.0001***		0.056
Group	1, 48	1.88	0.17		0.032
ConditionXGroup	1, 48	0.56	0.54		0.001
Posterior electrodes					
Condition	1, 48	17.58	0.0001***		0.064
Group	1, 48	4.07	0.04*		0.069
ConditionXGroup	1, 48	0.003	0.95		0.000013
450-to-600ms time window					
Condition	1, 48	4.91	0.03*		0.016
Group	1, 48	1.86	0.17		0.023
Longitude	2, 96	29.56	< 0.0001***	0.68	0.084
GroupXLongitude	2, 96	3.01	0.07	0.68	0.009
ConditionXGroup	1, 48	2.99	0.09		0.010
ConditionXLongitude	2, 96	8.84	0.001**	0.71	0.013
ConditionXGroupXLongitude	2, 96	0.33	0.64	0.71	0.0005
Group: AE - Cluster: Af3, Af4, Fp1, Fp2, Fpz					
Condition	1, 26	0.12	0.73		0.0007
Longitude	2, 52	8.93	0.003**	0.59	0.047
ConditionXLongitude	2, 52	8.020	0.003**	0.71	0.016
Group: N-AE					
Condition	1, 22	7.92	0.001**		0.058
Longitude	2, 44	24.07	<0.0001	0.67	0.14
ConditionXLongitude	2, 44	2.33	0.12	0.69	0.010

Cluster: Af3, Af4, Fp1, Fp2, Fpz					
Condition	1, 22	0.45	0.50		0.005
Cluster: C1, C2, Cp1, Cp2, Cpz, Cz, P1, P2, Pz					
Condition	1, 22	10.07	0.004**		0.11

Table VII - Multiple Regression Statistics of the final N400 model as resulting from the backward procedure and coefficients values.

N400 Model	**Start AIC**	**Final AIC**	**F**	**R^2**	**ΔR^2**	**p**		
	62.42	54.94	(5,44)=3.44	0.28	0.19	<0.05*		
Coefficients	**Estimate**	**Std. Error**	**t**	**p <**	**B_i**	**VIF**	**Mean VIF**	**Autocorrelation**
SS	0.71	0.28	2.54	0.05*	0.39	1.45		
Fluency	- 0.43	0.26	- 1.62	n.s.	- 0.23	1.29		
ART	1.23	0.35	3.29	0.005**	0.67	2.28		
Group	- 1.82	0.75	- 2.41	0.05*	- 0.50	2.63		
ART X Group	- 1.08	0.72	- 1.49	n.s.	- 0.59	1.40		
							1.81	D-W=2.27; p=n.s.

Conclusions

A thorny truth?

The main aim of this dissertation was to investigate the role of individual variation in metaphor interpretation and cast light on if and to what extent speakers' characteristics influence metaphor processing, an issue still rather unexplored within experimental research on metaphor. To this purpose, two specific issues were addressed. The first issue investigated the role of a set of cognitive abilities in the development of metaphor comprehension (Study 1). The second issue concerned individual variation in metaphor processing during adulthood and focused on the role of Education background and neuropsychological performance (Study 2).

Study 1 used a newly developed behavioural paradigm to tease apart the contribution of Alternative Naming and Analogy Perception in the development of metaphor understanding. Both Alternative Naming (i.e., the ability to understand and accept that one and the same entity can take two or more linguistic labels) and Analogy Perception (i.e., the ability to identify similarities across entities based on a shared relation) are essential to metaphor comprehension. However, they follow a different developmental trajectory and this may be crucial to explain metaphor development in pre-school years since difficulties in one of both of them may increase the cognitive costs required to access metaphorical meaning, hence impede or slow down young children's understanding of metaphor.

In this study, 3- and 4-year-olds were tested in three separate tasks which were administered within the same picture-matching general procedure and which assessed, respectively, children's ability to understand metaphorical versus literal language (i.e., the Metaphor task), their alternative naming abilities (i.e., the Alternative Naming Task), and their ability to perceive analogy (i.e., the Analogy Perception Task). Three main results emerged from Study 1. First, children of both ages were above chance in metaphor understanding but, importantly, they exhibited significantly more difficulties with metaphoric than literal language. Second, when not grasping the meaning of a metaphor, they interpreted it literally by literally interpreting both the metaphor target and vehicle. Third, and most importantly for the purpose of the study, pre-schoolers' ability to understand metaphor varied depending on their proficiency in alternative naming and analogy perception. Depending on their developmental trajectory, both Alternative Naming and Analogy Perception influence pre-schoolers' comprehension of metaphor and act as enhancing or impeding factors. At age 3, both skills seem to hinder metaphor understanding and, thus, are both likely to increase the associated cognitive demands. At age 4, children's difficulties with alternative naming are fully solved while this is not the case for analogy perception. At this stage of development, while alternative naming is likely to enhance metaphor comprehension, the ability to perceive analogy seems not developed enough yet to ease the associated cognitive effort and, as such, constitutes an impeding factor. Taken together, these results show first and foremost that metaphor comprehension develops as long as assigning two labels to the same entity and detecting relational similarities across entities develop too. Thus, individual variation in cognitive development plays a prominent role in the development of metaphor understanding and this can explain, at least partially, the developmental path for metaphor.

Study 2 investigated individual variation in adults' metaphor processing. In particular, two main factors likely to be responsible for the variation were targeted: first, a socio-demographical trait such as Education type; second, cognitive performance as measured by a set of neuropsychological tests which assessed for verbal working memory, exposure to reading, verbal fluency, IQ and autism quotient. Two groups of participants, with an academic and a non-academic education, read metaphorical and literal expressions while their EEG was recorded and, afterwards, were administered the battery of neuropsychological tests. Participants were balanced for age and gender, thus they solely differed in terms of the education programme they had attended. Study 2 revealed the following interesting results. First, the brain electric activity associated to metaphor processing - as compared to literal language processing - varied depending on education type. In fact, a different ERPs pattern was found in the two groups of participants. People with an academic education exhibited the standard N400-P600 biphasic pattern, respectively linked to lexical access shaped by context and pragmatic interpretation. Interestingly, people with a non-academic education showed a N400 component only, which prolonged over time until at least 600ms post-stimulus onset. As discussed in chapter 2, these differences are likely to mirror different cognitive costs for the two groups' understanding of metaphor. Alternatively, they might reveal different interpretive strategies for metaphor comprehension depending on education type (see § 4.1). However, what is crucial here is that the emerged ERPs differences first and foremost provide evidence for one main point: even in adulthood, when the cognitive system is presumably mature enough, metaphor processing is not the same for everyone. These data show that it is significantly influenced by the type of education an individual has received.

Second, a multiple regression analysis revealed that education background accounted for 50% of the variance in the N400 amplitude to metaphors, yet exposure to reading - as an indicator for individuals' semantic network - and verbal working memory also facilitated the lexical process reflected in the negativity. This pattern of results suggest that education background does affect metaphor processing, by modulating the cognitive efforts as indexed in the ERP components. However, variation in the brain response is also significantly explained by individual differences in cognitive performance. Both exposure to reading and verbal working memory contribute to metaphor processing: the more efficient are these abilities in an individual, the less the cognitive costs associated to the lexical access phase as indexed by the N400 effect, hence the easier metaphor processing.

In a word, then, both in development and adulthood, novel metaphor interpretation is associated to individual differences. Moreover, this variation seems to depend on at least two individual sources: endogenous sources, such as differences in cognitive functioning, as well as environmental sources, such as education. This overall finding raises some implications for research on metaphor comprehension and processing. More broadly, it also raises far from trivial implications for the architecture of the communicative mind. These two points are now addressed.

Implications for research on metaphor

As mentioned in the introduction, experimental research has shown that interpreting metaphorical language is costlier than interpreting literal language. In addition, these costs are not always the same but they vary depending on what the metaphor is like.

Novel/unfamiliar and literary metaphors involve extra-cognitive effort than conventional and familiar metaphors (e.g., Arzouan et al., 2007b; Bambini et al., 2018) and interpreting metaphor out of context requires more processing resources than when the interpretive process is facilitated by a supporting context (Bambini et al., 2016a). Consistently with this, young children have been reported to either not accessing metaphorical meaning or exhibiting remarkable difficulties and metaphor comprehension is traditionally thought to emerge fairly late in development (see Winner, 1988/1997). Overall, then, it seems that metaphor understanding is not a simple, effortless, 'unitary' process. Instead, it seems complex enough at least to (i) require a given expenditure of cognitive resources that individuals might not possess at all stages of the lifespan (i.e., development), and (ii) be largely subject to variation in terms of processing costs even during adulthood, depending on the type of metaphor and the communicative circumstances in which it occurs (i.e., with or without context). The data emerged in this dissertation provide an interesting contribution to the overall experimental purview. They suggest that understanding metaphor - and the amount of the associated cognitive costs - does not only depend on the linguistic features of the stimulus *per se*, but it also depends on the speakers' characteristics. In other words, understanding one and the same metaphor might not prove the same for everyone, but it may require (at least quantitatively) different cognitive resources depending precisely on differences among individuals. This dissertation has focused on specific sets of factors that relate to individual differences (i.e., alternative naming and analogy perception for development, cognitive performance and education type in adulthood). Further research would now be worth conducting to further explore this line of investigation and not only identifying other possible factors or clusters of factors, but also mapping their interactions with the linguistic features that are known to influence metaphor processing. For example, does education type equally influence the ERPs pattern for familiar and unfamiliar metaphor? How does context interact with education type? Do all types of metaphor follow the same developmental trajectory? If no, where do developmental differences across metaphors stem from?

The idea that metaphor interpretation is bound to individual variation emerges as a fairly reasonable conclusion from the findings of Study 1 and 2. In turn, this insinuates that metaphor understanding might not only be a matter of going beyond literalness by passing or not through literal meaning (cf., theoretical models on metaphor comprehension).

Rather, and plausibly, it might also be a matter of *how able an individual* is to handle the cognitive costs associated to metaphor interpretation, depending on endogenous as well as exogenous factors. This creates intriguing implications for theories of metaphor understanding. In fact, independently of the conspicuous differences between one theory and the other (e.g., Relevance Theory vs. Giora's Graded Salience Hypothesis), traditional theories of metaphor comprehension assume one and only one interpretive strategy and this is in turn developed from the assumption that the cognitive costs required by metaphorical language – if any – are the same independent of both the metaphor and the speakers' characteristics. Yet, in light of Study 1 and Study 2 - together with other relevant evidence (see introduction) - one very different scenario seems to knock on our door: individual differences might systematically influence metaphor comprehension and its underlying cognitive efforts. Whether or not this also leads or implies qualitative differences in the interpretive strategy involved remains an open question and certainly deserves attention with more purposely fine-tuned experimental work. However, the theoretical debate on metaphor has not addressed these issues up to now. Yet, empirical evidence already demonstrates associations between individual variables and metaphor interpretation. As a consequence, theories of metaphor comprehension should work towards a reassessment of the implicit assumption on which they are based because empirical evidence seems to suggest that the story of how metaphor is interpreted might reveal even more sophisticated than previously thought. It might indeed vary depending on *who you are*.

Implications for the architecture of the communicative mind

Studies 1 and 2 addressed individual variation in metaphor understanding by looking at two different stages in the lifespan, namely development and adulthood. Accordingly, several factors and cognitive functions were looked at in relation to metaphor interpretation, given that different research questions specifically targeted development and adulthood (i.e., Study 1: what are the 'basics' for metaphor understanding? Do pre-schoolers possess them in a way mature enough to deal with metaphor? Study 2: do education type and variation in cognitive performance influence adults' metaphor processing?). Importantly, however, both studies shows that novel metaphor understanding changes depending on *who the processor is*. The cognitive costs required to comprehend metaphor vary depending on individual differences. More essentially, this is the case for both development and adulthood. And, even more essentially, the studies - as a whole - bring to light one crunchy broad fact: several cognitive functions appear to be involved during metaphor interpretation.

Though fairly new within the literature on metaphor comprehension, the patterns above are not that surprising. In fact, they fit smoothly with experimental literature exploring individual differences in the language system. There is a consistent bulk of empirical work demonstrating that individual differences greatly influence most of the language sub-systems, not only in atypical population – notably, the traditional and paradigmatic stage for researching individual differences – but also across typical population. Starting from development, individual differences are large and notably stable across first language acquisition (Bates et al., 1995; Bornstein et al., 2018; Clark, 2009). They are also observed from very early on and across all linguistic domains. For example, variation in auditory brainstream responses in 6-week-old infants predicts their emerging language knowledge at 9 months (Chonchaiya et al., 2013). Family SES impacts the efficiency of language proficiency already at 24 months (Fernald et al., 2013) and, by the same age, the size of productive vocabulary enhances kids' prediction skills during language comprehension (Mani and Huettig, 2012). Individual differences, however, do remain in adulthood too. For instance, language proficiency models the adults' electrophysiological response to syntactic processing (Pakulak & Neville, 2010) and even a genetic factor such as familial sinistrality impacts the hemispheric sensitivity to adults' processing of syntactic violations (Lee & Federmeier, 2015). Consistently with this, semantic processing too is subject to individual variation. To mention a few examples, inhibitory abilities enhance adults' processing of the abstract relations between words (Boudewyn et al., 2012); people with a higher exposure to reading assign pronouns to the grammatical subject more consistently (Arnold et al., 2018); and vocabulary proficiency and formal literacy influence predictive and anticipatory mechanisms in sentence processing (Borovsky et al., 2012; Mishra et al., 2012). On top of this, recent studies have revealed sex and gender differences in (early) semantic processing (Xu et al., 2014; Wirth et al., 2007) and in the ERPs correlates to semantic priming (Daltrozzo et al., 2007).

Individual differences, then, seem pervasive in language and the results of studies 1 and 2 fit perfectly with this broader picture. Importantly, the existence of individual differences has already been acknowledged within the theoretical debate on the architecture of the linguistic system (see Kidd et al., 2017 for a good review on the topic). In fact, against the formal linguistic approaches (e.g., Lewis et al., 2006), a broader class of so-called emergentist approaches developed. Very briefly, the formal linguistic accounts do not leave much space to individual differences in language since, inspired by the Chomskyian theory of a Universal Grammar according to which the principles of grammar are universal and innate (Chomsky, 1965), they approach language as a vertical faculty, that is a module specialized for the processing of information for one single, specific domain such as language or faces. Conversely, the emergentist approaches (e.g., the usage-based approach to language acquisition and experience, Lieven, 2016; the constraint-based approaches to language processing, Bates et al., 1989; MacDonald & Christiansen, 2002) eschew this rigid idea of an innately specified knowledge (e.g., Universal Grammar). Instead, they conceive of linguistic ability as a horizontal faculty, that is a system that processes information across several domains (e.g., attention, memory, executive functions) and argue that language is learnable via several processes that may not necessarily be specific to language. As such, the emergentist accounts predict patterns of interactions and associations between the linguistic system and both the language and other cognitive systems as well as other environmental factors originating from usage and experience (see Kidd et al., 2017). In sum, individual differences in the language system represent a solid reality within the correspondent theoretical debate and attempts have been made to map this variation into tenable models of the linguistic mind. Hence, the results of Studies 1 and 2 fits well with both experimental and theoretical research on individual variation in linguistics. What about pragmatics?

As mentioned throughout this dissertation, there is not much experimental research that has directly targeted individual differences in metaphor processing and development. In this respect, the patterns of results revealed by Studies 1 and 2 provide an insightful contribution to the scientific debate. However, these results do not stay out of the blue in the broader purview of research in pragmatic processing. Rather, they fit smoothly with a few recent experimental studies reporting individual differences in other pragmatic phenomena. For example, in a recent EEG experiment on humour processing (Canat et al., forthcoming), we have found that AQ correlates with the size of the Left Anterior Negativity, an ERP component indexing the earliest detection of incongruity during humour processing and suggesting that social skills affect humour comprehension in the earliest processing phase. In another recent study, Domaneschi & Di Paola (2019) found that verbal working memory enhances presupposition processing in young adults and that the ability to recover presupposed information is affected by normal aging. This series of studies is fully consistent with recent experimental literature on the topic. In fact, it has been shown that individual variation in cognitive resources and AQ modulates scalar implicatures processing both in terms of the associated ERP response (Niewland et al., 2010) and behaviourally (Yang et al., 2018). Executive functions, especially inhibitory control, enhance people's perspective taking and can account for the (in)sensitivity to common ground information (Brown-Schmidt, 2009). Verbal ability and working memory predict the listeners' sensitivity to referential ambiguity by increasing the associated electrophysiological activity (Boudewyn et al., 2015). Moreover, individuals with high empathizing skills can rapidly integrate information about the speaker with the content of the message, exhibiting larger N400 effects to socially relevant information (van den Brink et al., 2012). As for development, evidence is even scanter than that for adulthood and most of this is the result of correlational rather than experimental research. Nonetheless, the general pattern as outlined above seems to hold in development too (see Matthews et al., 2018 for the only recent review on the topic). Medium to large associations are found between pragmatics and formal language (Bernard & Deleau, 2007; De Rosnay et al., 2014) and additional associations are found between mentalizing and irony understanding. For example, Filippova et al., (2008) found that both Theory of Mind (ToM) and language abilities make unique contribution to children's interpretation of irony, over and above age, memory and emotions detection. Also, two studies revealed that executive functions contribute more than IQ to a range of pragmatic skills in pre-school age (Blain-Brière et al., 2014) and that cognitive flexibility proves essential to the development of ambiguity detection in pre-schoolers (Gillis & Nilsen, 2014).

In sum, especially for typical population, pragmatic research on individual differences is still uneven and only a few studies have been conducted yet. Promisingly, however, the available data seem to point already towards a homogeneous direction. One in which – consistently with research on other linguistic areas (see above) – individual differences might end up representing the norm rather than the exception.

What stated above raises some potentially thorny issues about the mechanistic architecture of the pragmatic mind. Drawing from the Fodorian model of the mind as based on non-modular central systems (e.g., keeping and updating the represententations of the world, making decisions and plans, speculative and imaginative thinking) and modular peripheral systems (e.g., processing perceputual inputs such as linguistic inputs) (Fodor, 1993), pragmatics has been traditionally conceived of as a modular cognitive system, that is an autonomous component of the mind. Outstanding defensors of such a cognitive mentalist approach are notably Sperber & Wilson (1996; 2002). This approach sees pragmatics as an activity of intentions attribution and recognition. More in detail, pragmatic interpretation is a process in which the hearer infers the speaker's intended meaning from evidence she had provided for this purpose (i.e., utterances). On this view, pragmatics is argued to be a mental module, specialized in one single specific domain: mind-reading. More precisely, pragmatics is theorized to be a sub-module of the Theory of Mind module, that is a dedicated comprehension module with its own special principles and mechanisms (i.e., a module with dedicated processes for pragmatic interpretation as the one proposed by Relevance Theory; Sperber & Wilson, 1986/95).

However, Studies 1 and 2, together with the weight of experimental evidence reported above, pose some thorns to a modular view of pragmatics. For at least three reasons. First, but keep in mind that the standard modular view does not make any clear prediction, one should reasonably expect an involvement of ToM in the data across pragmatic phenomena. Yet, what emerges as a whole from experimental evidence is that multiple cognitive functions are involved in pragmatic processing, not just ToM. Second, even admitting that the pragmatics module might interact with other mental modules and that this could explain the co-involvement of several cognitive functions (but, again, notice that this is not addressed by the standard modular view), still, if pragmatics is an autonomous mind-reading sub-module, then one should expect a massive involvement of ToM, in all pragmatic phenomena and over and above other cognitive functions. Yet, though still restricted, the data seem to point towards a different direction. One that does not provide evidence in favour of a massive involvement of ToM nor it shows that ToM is "universally" involved in all pragmatic phenomena. Rather, on a closer look at the data, it might be the case that executive functions and language proficiency correlates with pragmatic processing more often than ToM[6]. Third, the essential point of any modular view of the mind, including the modular approach to pragmatics, is the existence of modules that are domain-specialized and only minimally interact to each other (see Mazzone, 2015). To put it simple, a mental module is by its own nature a rigid system. This is obviously difficult to reconcile with the variation in pragmatic processing revealed by the experimental literature since that variation relies precisely on several associations with several cognitive systems and environmental factors.

Attempts to reconcile the standard modular view of the mind with the flexibility of the mind itself have been recently made by Sperber (2005), who has proposed what he calls *massive modularity.* According to Sperber (2005), the organization of the cognitive processes into modules remains a pervasive phenomenon that pervasively rules mental activity. However, in this proposal, the mind is not seen anymore as working on the basis of central and peripheral systems (i.e., non-modular and modular, respectively; cf., the standard modular approach). Rather, there would only be modules which compete against each other. The flexibility of the mind would then be allowed by the competition/interaction among cognitive modules. Human behaviours - in the most cognitive sense of outcomes for a given cognitive activity - would then originate from a process of selection of contextually relevant information that results from the interactions between modules. Hence, the position of a massive modularity adopts a weaker definition of mental modules because it does not maintain anymore their original rigidity. Rather, it admits an interaction between one module and the other. This view might reconcile with data on individual differences in pragmatic processing, at least it might do so more than the standard modular account. In fact, it opens to the possibility of more than one cognitive system (i.e., module) involved for mental activity and this might have the potential to explain individual differences in the pragmatic data. However, the massive modularity thesis neither has spelled out yet any prediction about the co-occurrence of cognitive systems other than pragmatics nor it has somehow included individual variables (and the associated evidence) in the theoretical discussion. Furthermore, Sperber (2005) provides a new and weaker view of the modularity of mind in general. Issues on how this could apply to the

[6] The issue of whether or not pragmatics is all about intentions reading and ToM is out of the scope of this discussion. However, it is worth noticing that this issue is currently under debate and some authors are even adopting an extreme position, arguing that pragmatics and ToM do not completely overlap (Bosco et al., 2018).

pragmatic system in particular or what changes should be made to the standard view of pragmatics as a module have not been addressed yet. Thus, it remains unclear how the pragmatic module would work in light of massive modularity. Most relevant for the present discussion, it remains unclear how any (stronger or weaker) modular account could explain individual differences in pragmatics - that are found indeed, in development as well as adulthood.

Beyond this, if on the one hand a weaker notion of modularity might somehow fit the experimental data, on the other hand one might sensibly wonder whether at this point it is still sensible to theoretically approach pragmatics in modular terms. That is to say, once the mind is admitted a more flexible mechanistic architecture (i.e., modules interact with each other and behaviours are the outcomes of these interactions) - and this seems confirmed by the data on individual differences - is it still appropriate to speak of modularity and, ultimately, of pragmatics as an autonomous component? Why then not to commit to more holistic approaches to the mind? Holistic approaches propose an account of mental functioning as based on the notion of *network*. For example, Fuster (2003) claims that the main higher-order cognitive functions work holistically, that is there is no cortical area or network that is exclusively dedicated to each single cognitive function. Rather, perception, memory, attention, intelligence *and language* are argued to share the same underlying neural cells and connections. Now, it is out of my aim to discuss exaustively the two positions above (i.e., (massive) modularity and cognitive holism) and a detailed discussion in this respect can be found in Mazzone (2015) (see also Mazzone, 2016 for discussion on the (un)tenability of the modularity thesis). The main point here is that the results of Studies 1 and 2, together with increasing experimental evidence, suggest that pragmatic ability seems to involve processes that engage several cognitive systems. This clearly brings us far away from an idea of pragmatics as an entirely autonomous component, rather it brings to light that pragmatics too - as well as other cognitive systems- appears to be flexible. In light of this, then, the standard notion of a pragmatic module does not seem to render full justice to such a flexibility. Massive modularity might. But then, it becomes intuitively unclear how a weaker modularity account would differ from other more holistic accounts of cognition - at least in their substantial points. Based on the findings of Study 1 and 2 and on the rest of the experimental demonstrations, it is my suspect that - in the best of the possibilities - the existence of individual differences in pragmatic processing might represent a thorny issue for a conception of pragmatics as an autonomous component. Yet, the conception of pragmatics as an autonomous component is, currently, the only one (officially) on the market.

To conclude, the experiments conducted in this dissertation have shed light on the processing and comprehension of metaphor. They have shown that individual variation does characterize metaphor interpretation in both development and adulthood. This finding is in line with a recent bulk of experimental literature in pragmatics evidencing individual differences in a range of pragmatic phenomena. Taken together, the findings of Study 1 and 2 raise important implications to theories of metaphor interpretation and open a scenario that calls for the need to reconcile current theories of metaphor with the variation in the data. Last but not least, the overally emerged picture suggests that individual differences in pragmatic processing do exist. This poses not trivial implications for a view of pragmatics as an autonomous component of the mind. In fact, rather than representing error variance in the statistical models (as it has been the case for decades of experimental research), individual variables in pragmatic interpretation actually might represent a concrete, though thorny, truth.

References

Asch, S. & Nerlove, H. (1960). The development of double function terms in children: An exploratory investigation. In B. Kaplan & S. Wapner (Eds), *Perspectives in Psychological Theory: Essays in Honor of Heinz Werner.* New York, NY: International Universities Press, 47 – 60.

Acheson, D. J., Wells, J. B., & MacDonald, M. C. (2008). New and updated tests of print exposure and reading abilities in college students. *Behavior Research Methods*, 40, 278 – 289.

Aristotle (350 BCEa). *Poetics*. Translated by Butcher S. H. The internet classic archive <http://classics.mit.edu/Aristotle/poetics.1.1.html>.

Aristotle (350 BCEb). *Rhetoric*. Translated by Roberts Rhys W. The internet classic archive http://classics.mit.edu/Aristotle/rhetoric. html.

Arnold, J.E., Strangmann, I.M., Hwang, H., Zerkle, S., Nappa, R. (2018). Linguistic experience affects pronoun interpretation. *Journal of Memory and Language*, 102, 41 – 54.

Arzouan, Y., A. Goldstein and M. Faust. (2007a). Dynamics of hemispheric activity during metaphor comprehension: Electrophysiological measures. *Neuroimage*, 36. 222 – 231.

Arzouan, Y., A. Goldstein and M. Faust. (2007b). Brainwaves are stethoscopes: ERP correlates of novel metaphor comprehension. *Brain Research*, 1160, 69 – 81.

Baayen, R.H., Davidson, D.J., Bates, D.M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59, 390 – 412.

Bambini, V., Gentili, C., Ricciardi, E., Bertinetto, P. M., & Pietrini, P. (2011). Decomposing metaphor processing at the cognitive and neural level through functional magnetic resonance imaging. *Brain Research Bulletin*, 86, 203 – 216.

Bambini, V. & Resta, D. (2012). Metaphor and experimental pragmatics: When theory meets empirical investigation. *HumanaMente Journal of Philosophical Studies*, 23, 37 – 60.

Bambini, V., Ghio, M., Moro, A., & Schumacher, P. B. (2013). Differentiating among pragmatic uses of words through timed sensicality judgments. *Frontiers in Psychology*, 4, 938.

Bambini, V., Bertini, C., Schaeken, W., Stella, A., & Di Russo, F. (2016a). Disentangling metaphor from context: An ERP study. *Frontiers in Psychology*, 7, 559.

Bambini V., Arcara, G., Bechi, M., Buonocore, M., Cavallaro, R., Bosia, M. (2016b). The communicative impairment as a core feature of schizophrenia: Frequency of pragmatic deficit, cognitive substrates, and relation with quality of life. *Comprehensive Psychiatry*, 71, 106 – 120.

Bambini, V., Canal, P., Resta, D. & Grimaldi, M. (2018). Time Course and Neurophysiological Underpinnings of Metaphor in Literary Context. *Discourse Processes.*

Baron- Cohen, S. (1995). *Mindblindness: an essay on autism and theory of mind*. Cambridge, MA: MIT Press.

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The autism-spectrum quotient (AQ): Evidence from asperger syndrome/high-functioning autism, malesand females, scientists and mathematicians. *Journal of Autism and Developmental Disorders*, 31, 5 – 17.

Barr, D.J., Levy, R., Scheepers, C., Tily, H.J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language,* 68, 255 – 278.

Bates, D., Mächler, M., Bolker, B.,Walker, S., Christensen, R. H. B., Singmann, H., et al. (2015). L. Package 'lme4'. *Convergence*, 12, 1.

Bates, E., Bretherton, I., Snyder, L., Beeghly, M., Shore, C., McNew, S., Carlson, V., Williamson, C., Garrison, A., et al. (1988). *From first words to grammar: Individual differences and dissociable mechanisms.* New York, NY, US: Cambridge University Press.

Berger, F., & Höhle, B. (2012). Restrictions on addition: Children's interpretation of the focus particles auch ('also') and nur ('only') in German. *Journal of Child Language*, 39, 383 – 410.

Bernard, S., Deleau, M. (2007). Conversational perspective-taking and false belief attribution: A longitudinal study. *British Journal of Developmental Psychology*, 25, 443 – 460.

Billow, R. M. (1975). A cognitive developmental study of metaphor comprehension. *Developmental Psychology*, 11, 415.

Black, M.(1962) *Models and Metaphors*. Ithaca: Cornell University Press.

Blain-Brière, B., Bouchard, C., Bigras, N. (2014). The role of executive functions in the pragmatic skills of children age 4 -5. *Frontiers in Psychology*, 5, 240.

Blasko, D. G. (1999). Only the tip of the iceberg: Who understands what about metaphor? *Journal of Pragmatics*, 31, 1675 – 1683.

Bohrn, I. C., Altmann, U., and Jacobs, A. M. (2012). Looking at the brains behind figurative language - A quantitative meta-analysis of neuroimaging studies on metaphor, idiom, and irony processing. *Neuropsychologia*, 50, 2669 – 2683.

Bornstein, M., & Putnick, D.L. (2012) Stability of language in childhood: a multiage, multidomain, multimeasure, and multisource study. *Developmental Psychology*, 48, 477 – 491.

Bosco, F., Tirassa, M., Gabbatore, I. (2018). Why Pragmatics and Theory of Mind Do Not (Completely) Overlap. *Frontiers in Psychology*, 9, 1453.

Boudewyn, M.A:, Long, D.L., Traxler, M. J., Lesh, T.A., Dave, S., et al. (2015). Sensitivity to Referential Ambiguity in Discourse: The Role of Attention, Working Memory, and Verbal Ability. *Journal of Cognitive Neuroscience*, 27, 2309 – 2323.

Boudewyn, M.A., Long., D.L., Swaab, T. (2012). Cognitive Control influences the use of meaning relations during spoken sentence comprehension. *Neuropsychologia*, 50, 2659 – 2668.

Bowdle, B. & Gentner, D. (2005). The career of metaphor. *Psychological Review*, 112, 193 – 216.

Brouwer, H., Fitz, H., & Hoeks, J. (2012). Getting real about Semantic Illusions: Rethinking the functional role of the P600 in language comprehension. *Brain Research*, 1446, 127 – 143.

Brown-Schmidt, S. (2009). The role of executive function in perspective taking during online language comprehension. *Psychonomic Bulletin & Review*, 16, 893 – 900.

Burkett, C. & Goldman, S.R. (2016). "Getting the Point" of Literature: Relations between Processing and Interpretation. *Discourse Processes*, 53, 457 – 487.

Canal, P., Bischetti, L., Di Paola, S., Ricci, I., Bertini, C., Bambini, V. (forthcoming). 'Honey, shall I change the baby? - Well done, choose another one': ERP and Time-Frequency correlates of humor processing. Accepted for pubblication in *Brain and Cognition*.

Canal, P., Pesciarelli, F., Vespignani, F., Molinaro, N., & Cacciari, C. (2017). Basic Composition and Enriched Integration in Idiom Processing: An EEG Study. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 43, 928 – 943.

Carlson, S.M & Moses, L.J. (2001). Individual Differences in Inhibitory Control and Children's Theory of Mind. *Child Development*, 72, 1032 – 1053.

Carriedo N., Corral A., Montoro, P.R., Herrero, L., Ballestrino, P., Sebastián, I. (2016). The Development of Metaphor Comprehension and Its Relationship with Relational Verbal Reasoning and Executive Function. *PLoS ONE* ,11, e0150289.

Carston, R. (2002). *Thoughts and Utterances*. Blackwell Publishing.

Carston, R. (2010). XIII-Metaphor: Ad Hoc Concepts, Literal Meaning and Mental Images. *Proceedings of the Aristotelian Society* (Hardback), 110, 295 – 321.

Carston, R., & Wearing, C. (2011). Metaphor, hyperbole and simile: A pragmatic approach. *Language and Cognition*, 3, 283 – 312.

Caselli M.C. & Casadio, P. (1995). *Il Primo Vocabolario del Bambino. Guida all'uso del questionario MacArthur per la valutazione della comunicazione e della comunicazione e del linguaggio nei primi anni di vita*. Milano: Franco Angeli.

Cassirer, E. (1946). *Language and Myth*. New York: Dover Publication Inc.

Champagne-Lavau, M., Monetta, L., Moreau, N. (2012). Impact of Educational Level on Metaphor Processing in Older Adults. *Revue française de linguistique appliquée*, 12, 89 – 100.

Chen, Z., Sanchez, R.P., & Campbell, T. (1997). From beyond to within their grasp: Analogical problem solving in 10- and 13-month-olds. *Developmental Psychology*, 33, 790 – 801.

Chiappe, D. & Chiappe, P. (2007). The role of working memory in metaphor production and comprehension. *Journal of memory and language*, 56, 172 – 188.

Chomsky, N. (1965). *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.

Chonchaiya, W., Tardif, T., Mai., X, Xu, L., Li, M., et al. (2013) Developmental trends in auditory processing can provide early predictions of language acquisition in young infants. *Developmental Science*, 16, 159 – 172.

Clark, E.V. (1997). Conceptual perspective and lexical choice in acquisition. *Cognition*, 64, 1 – 37.

Clark, E.V. (2009). *First Language Acquisition*. Cambridge: Cambridge University Press.

Columbus, G., Sheikh, N. A., Côté-Lecaldare, M., Häuser, K., Baum, S. R., & Titone, D. (2015). Individual differences in executive control relate to metaphor processing: an eye movement study of sentence reading. *Frontiers in Human Neuroscience*, 8, 1057.

Cometa, M., & Eson, M. (1978). Logical operations and metaphor interpretation: A piagetian model. *Child Development*, 49, 649 – 659.

Coulson, S., & Van Petten, C. (2002). Conceptual integration and metaphor: An event-related potential study. *Memory & Cognition*, 30, 958 – 968.

Coulson, S. & van Petten, C. (2007). A special role for the right hemisphere in metaphor comprehension? ERP evidence from hemifield presentation. *Brain Research*, 1146, 128 – 145.

Daltrozzo, J., Wioland, N., Kotchoubey, B. (2007). Sex Differences in Two Event-Related Potentials Components Related to Semantic Priming. *Archives of Sexual Behavior*, 36, 555 – 568.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, 19, 450 – 466.

De Grauwe, S., Swain, A., Holcomb, P. J., Ditman, T., & Kuperberg, G. R. (2010). Electrophysiological insights into the processing of nominal metaphors. *Neuropsychologia*, 48, 1965 – 1984.

De Rosnay, M., Fink, E., Begeer, S., Slaughter, V., Peterson, C. (2014). Talking theory of mind talk: Young school- aged children's everyday conversation and understanding of mind and emotion. *Journal of Child Language*, 41, 1179 – 1193.

Deamer, F. & Pouscoulous, N. (submitted). Non-literal children: Pre-schoolers' understanding of figurative language.

Declercq, C., Baltazart, V., Didon, D. (2010). Les nuages sont un troupeau de moutons : a` quel aˆge les enfants comprennent-ils qu'il n'y pas de moutons dans le ciel? *Canadian Journal of Experimental Psychology*, 64, 142 – 152.

Doherty, M. J., & Perner, J. (1998). Metalinguistic awareness and Theory of Mind: Just two words for the same thing? *Cognitive Development*, 13, 279 – 305.

Domaneschi, F. & Di Paola, S. (2019). The aging factor in presupposition processing. *Journal of Pragmatics*, 140, 70 – 87.

Falkum, I.L., Recasens, M., Clark, E. V. (2017). “The moustache sits down first”: on the acquisition of metonymy. *Journal of Child Language*, 44, 87 – 119.

Federmeier, K.D., Wlotko, E.W., De Ochoa-Dewald, E., Kutas, M. (2007). Multiple effects of sentential constraint on word processing. *Brain Research*, 1146, 75 – 84.

Federmeier, K.D, Kutas, M., Schul, R. (2010). Age-related and individual differences in the use of prediction during language comprehension. *Brain and Language*, 115, 149 – 161.

Fernald, A., Marchman, V.A., Weisleder, A. (2013). SES differences in language processing skill and vocabulary are evident at 18 months. *Developmental Science*, 16, 234 – 248.

Filippova, E., Astington, J. (2008). Further development in social reasoning revealed in discourse irony under- standing. *Child Development*, 79, 126 – 138.

Fodor, J. (1983) *Modularity of Mind*. Cambridge, MA: MIT Press.

Fuster, J. (2003). *Cortex and Mind*. Oxford: Oxford University Press.

Gentner, D. (1977). If a tree had a knee, where would it be? Children’s performance on simple spatial metaphors. *Papers and Reports on Child Language Development*, 13, 157 – 164.

Gentner, D. (1988). Structure-mapping in analogical development: The relational shift. *Child Development*, 59, 47 – 59.

Gentner, D. (1989). The Mechanisms of Analogical Learning. In S. Vosniadou & A. Ortony (Eds.) *Similarity and Analogical Reasoning*. London: Cambridge University Press, 199 – 241.

Gentner, D., Bowdle, B. F., Wolff, P., & Boronat, C. (2001). Metaphor is like analogy. In D. Gentner, K. J. Holyoak, & B. N. Kokinov (Eds.), *The Analogical Mind: Perspectives from Cognitive Science*. Cambridge, MA: MIT Press, 199 – 253.

Gernsbacher , A.M. (1993). Less Skilled Readers have less efficient suppression mechanims. *Psychological Science*, 4, 294 – 298.

Gibbs, R. W. (1994). *The Poetics of Mind: Figurative Thought, Language, and Understanding.* Cambridge: Cambridge University Press.

Gibbs, R.W. Jr. (2010). The Dynamic Complexities of Metaphor Interpretation. *D.E.L.T.A*, 26:especial, 657 – 677.

Gibbs, R.W. Jr. & Colston, H.L. (2012). *Interpreting Figurative Meaning.* NY: Cambridge University Press.

Gildea, P. & Glucksberg, S. (1983). On Understanding Metaphor: The Role of Context. *Journal of verbal learning and verbal behaviour*, 22, 577 – 590.

Gillis, R., & Nilsen, E. (2014). Cognitive flexibility supports preschoolers' detection of communicative ambiguity. *First Language*, 34, 58 – 71.

Giora, R. (1997). Understanding figurative and literal language: the graded salience hypothesis. *Cognitive Linguistics*, 8, 183 – 206.

Giora, R. (2008). Is metaphor unique? in R.W. Gibbs (Ed.) *The Cambridge Handbook of Metaphor and Thought*. Cambridge: Cambridge University Press, 143 – 160.

Glucksberg, S., Gildea, P., Bookin, H.B. (1982). On Understanding Nonliteral Speech: Can People Ignore Metaphors? *Journal of Verbal Learning and Verbal Behaviour* , 21, 85 – 98.

Glucksberg, S. (2001). *Understanding Figurative Language: From Metaphor to Idioms*. Oxford: Oxford University Press.

Glucksberg, S., Newsome, M.R., Goldvarg, Y. (2001). Inhibition of the Literal: Filtering Metaphor-Irrelevant Information During Metaphor Comprehension. *Metaphor & Symbol*, 16, 277 – 298.

Glucksberg, S. (2003). The psycholinguistics of metaphor. *Trends in Cognitive Sciences*, 7, 92 – 96.

Glucksberg, S. (2008). How metaphors create categories - quickly. In R.W. Gibbs (ed.) *The Cambridge Handbook of Metaphor and Thought.* Cambridge: Cambridge University Press, 67 – 83.

Goldstein, A., Y. Arzouan and M. Faust. (2012). Killing a novel metaphor and reviving a dead one: ERP correlates of metaphor conventionalization. *Brain and Language*, 123, 137 – 142.

Goswami, U. (1989). Relational Complexity and the development of analogical reasoning. *Cognitive Development*, 4, 251 – 268.

Goswami, U., & Brown, A.L. (1989). Melting chocolate and melting snowmen: Analogical reasoning and causal relations. *Cognition*, 35, 69 – 95.

Goswami, U. & Brown, A.L. (1990). Higher-order structure and relational reasoning: Contrasting analogical and thematic relations. *Cognition*, 36, 207 – 226.

Goswami, U. (1991). Analogical reasoning: What develops? A review of research and theory. *Child Development*, 62, 1 – 22.

Goswami, U. (2001). Analogical reasoning in children. In K. Holyoak, D. Gentner, & B. Kokinov (Eds), *The Analogical Mind: Perspectives from Cognitive Science*. Cambridge, MA: The MIT Press, 437 – 470.

Grassmann, S., & Tomasello, M. (2009). Young children follow pointing over words in interpreting acts of reference. *Developmental Science*, 13, 252 – 263.

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds.), *Syntax and Semantics: Speech Acts* (Vol. 3). New York, NY: Academic Press, 41 – 58. Reprinted in Grice (1989), 22 – 40.

Grice, H. P. (1989). *Studies in the Ways of Words*. Cambridge, MA: Harvard University Press.

Groppe, M.D., Makeig, S., Kutas, M. (2008). Independent Component Analysis of Event-Related Potentials. *Cognitive Science Online*, 6.1, 1 – 44.

Grosse, G., Schulze, C., Noveck, I,, Tomasello, M. & Katsos, N. (2011). Three-year-olds make inferences based on informativeness. *Paper presented at the 4th Biennial Conference of Experimental Pragmatics*, June 2-4 2011, Universitat Pompeu Fabra, Barcelona, Spain.

Hagoort, P., C. Brown and J. Groothusen. (1993). The syntactic positive shift (SPS) as an ERP measure of syntactic processing. *Language and Cognitive Processes*, 8, 439 – 483.

Huff, F. J., Corkin, S., & Growdon, J. H. (1986). Semantic impairment and anomia in Alzheimer’s disease. *Brain and Language*, 28, 235 – 249.

Katsos, N. & Bishop, D. V. M. (2011). Pragmatic tolerance: Implications for the acquisition of informativeness and implicature. *Cognition*, 120, 67 – 81.

Kazmerski, V. A., Blasko, D. G., & Dessalegn, B. G. (2003). ERP and behavioral evidence of individual differences in metaphor comprehension. *Memory & Cognition*, 31, 673 – 689.

Kidd, E.; Donnelly, S., Christansen, M.H. (2017). Individual Differences in Language Acquisition and Processing. *Trends in Cognitive Sciences*, 22, 154 – 169.

Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What’s new in Psychtoolbox-3. *Perception*, 36, 1 – 16.

Kolk, H.H.J., Chwilla, D.J., van Herten M., & Oor, P.J.W. (2003). Structure and limited capacity in verbal working memory: A study with event-related potentials. *Brain and Language*, 85, 1 – 36.

Kutas, M. and S.A. Hillyard. (1980). Event-related brain potentials to semantically inappropriate and surprisingly large words. *Biological Psychology*, 11, 99 – 116.

Kutas, M. & Federmeier, K.D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event related brain potential (ERP). *Annual Review of Psychology*, 62, 621 – 647.

Lai, V. T., Curran, T., & Menn, L. (2009). Comprehending conventional and novel metaphors : An ERP study. *Brain Research*, 1284, 145 – 155.

Lakoff, G. & Johnsen, M. (1980). *Metaphors We Live By*. Chicago, IL: The University of Chicago Press.

Lakoff, G. & Johnsen, M. (2003). *Metaphors we live by*. 2nd edn. London: The university of Chicago press.

Lee, C. and Federmeier, K.D. (2011). Differential age effects on lexical ambiguity resolution mechanisms. *Psychophysiology*, 48, 960 – 972.

Lee, C., Federmeier, K.D. (2015). It's All in the Family: Brain Asymmetry and Syntactic Processing of Word Class. *Psychological Science*, 26, 997 – 1005.

Lewandowsky, S., Oberauer, K., Yang, L.-X., & Ecker, U. K. (2010). A working memory test battery for MATLAB. *Behavior Research Methods*, 42, 571 – 585.

Lewis, R.L., Vasishth, S., Van Dyke, J.A. (2006). Computational principles of working memory in sentence comprehension. *Trends in Cognitive Sciences*, 10, 447 – 454.

Lieven, E.V.M. (2016). Usage-based approaches to language development: where do we go from here? *Language and Cognition*, 8, 346 – 368.

MacDonald, M.C., Christiansen, M.H. (2002) Reassessing working memory: a comment on Just and Carpenter (1992) and Waters and Caplan (1996). *Psycholpgical Review*,109, 35 – 54 .

Mani, N., Huettig, F. (2012). Prediction During Language Processing is a Piece of Cake - But Only for Skilled Producers. *Journal of Experimental Psychology*, 38, 843 – 847.

Mashal, N., Vishne, T., Laor, N., Titone, D. (2013). Enhanced left frontal involvement during novel metaphor comprehension in schizophrenia: evidence from functional neuroimaging. *Brain and Language*, 124, 66 – 74.

Matthews, D., Lieven, E., & Tomasello, M. (2010).What's in a manner of speaking? Children's sensitivity to partner-specific referential precedents. *Developmental Psychology*, 46, 749 – 760.

Matthews, D., Biney, H., Abbot-Smith, K. (2018). Individual differences in children's pragmatic ability: A review of associations with formal language, social cognition and executive functions. *Language Learning and Development*, 14, 1 – 38.

Mazzone, M. (2015). *Razionalità pratica e linguaggio*. Roma: Corisco Edizioni.

Mazzone, M. (2016). Trivializing modularity. An associative-representational account of cognition. *Epistemologia*, 2, 201 – 215.

McElree, B., & Nordlie, J. (1999). Literal and figurative interpretations are computed in parallel. *Psychonomic Bulletin Review*, 6, 486 – 494.

Mennes, M., Wouters, H., Vanrumste, B., Lagae, L., & Stiers, P. (2010). Validation of ICA as a tool to remove eye movement artifacts from EEG/ERP. *Psychophysiology*, 47, 1142 – 1150.

Mishra, R.K., Singh, N., Pandey, A. (2012). Spoken language-mediated anticipatory eye movements are modulated by reading ability - Evidence from Indian low and high literates. *Journal of Eye Movement Research*, 5, 1 – 10.

Morriseau, T., Davies, C., Matthews, D. (2013). How do 3- and 5-year-olds respond to under- and over-informative utterances? *Journal of Pragmatics*, 59, 26 – 39.

Morrison, R.G., Doumas, L.A.A., Lindsey, E., Richland, L.A. (2011). A computational account of children's analogical reasoning: balancing inhibitory control in working memory and relational representation. *Developmental Science*, 14, 516 - 529.

Nelson, H. E. (1982). National Adult Reading Test (NART): For the assessment of premorbid intelligence in patients with dementia: Test manual, NFER-Nelson, Windsor, Uk.

Niewland, M., Ditman, T., Kuperberg, G.R. (2010). On the incrementality of pragmatic processing: An ERP investigation of informativeness and pragmatic abilities. *Journal of Memory and Language*, 63, 324 – 346.

Nippold, M. A. (1988/1998). *Later Language Development: The School-Age and Adolescent Years* (2nd Ed.). Austin, TX: Pro-Ed.

Norenzayan, A. & Heine, S.J. (2005). Psychological Universals: What Are They and How Can We Know? *Psychological Bulletin*, 131, 763 – 784.

Noveck, I, Bianco, M., Castri, A. (2001). The costs and benefits of metaphor. *Metaphor & Symbol*, 16, 79 – 91.

OECD (2012). Education at a Glance 2012: OECD Indicators, OECD Publishing.

Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia*, 9, 97 – 113.

Olkoniemi, H., Ranta, H., Kaakinen, J.K. (2016). Individual differences in the processing of written sarcasm and metaphor: Evidence from eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 42, 433 – 450.

Özçaliskan, S. (2005). On learning to draw the distinction between physical and metaphorical motion: Is metaphor an early emerging cognitive and linguistic capacity? *Journal of Child Language*, 32, 291 – 318.

Pakulak, E., & Neville, H. J. (2010). Proficiency differences in syntactic processing of monolingual native speakers indexed by event-related potentials. *Journal of Cognitive Neuroscience*, 22, 2728 – 2744.

Pearson, B. (1990). The comprehension of metaphor by preschool children. *Journal of Child Language*, 17, 185 – 203.

Perfetti, C. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading,* 11, 357 – 383.

Perner, J., Stummer, S., Sprung, M., & Doherty, M. (2002). Theory of Mind finds its Piagetian perspective: Why alternative naming comes with understanding belief. *Cognitive Development*, 103, 1 – 22.

Perner, J., Brandl, J.L., Garnham, A. (2003). What is a Perspective Problem? Developmental Issues in Belief Ascription and Dual Identity. *Facta Philosophica*, 5, 355 – 378.

Perner, J., Rendl, B., & Garnham, A. (2007). Objects of desire, thought, and reality: Problems of anchoring discourse referents in development. *Mind and Language*, 22, 475 – 513.

Perner, J. & Leahy, B. (2016). Mental Files in Development: Dual Naming, False Belief, Identity and Intensionality. *Review of Philosophy and Psychology*, 7, 491 – 508.

Peskin, J. (2010). The Development of Poetic Literacy During the School Years. *Discourse Processes,* 47(2), 77 – 103.

Petersen, I.T., Hoyniak., C.P., McQuillan, E.M., Bates, E.J., Staples, A.D. (2016) Measuring the development of inhibitory control: The challenge of heterotypic continuity. *Developmental Review*, 40, 25 – 71.

Pierce, R.S., MacLaren, R., Chiappe, D.L. (2010). The role of working memory in the metaphor interference effect. *Psychonomic Bulletin & Review*, 17, 400 – 404.

Pollio, H.R., Barlow, J.M., Fine, H.J. & Pollio, M.R. (1977). *Psychology and the Poetics of Growth: Figurative Language in Psychology, Psychotherapy, and Education*. Hillsdale, NJ: L Erlbaum.

Pouscoulous, N., Noveck, I., Politzer, G., & Bastide, A. (2007). Processing costs and implicature development. *Language Acquisition*, 14, 347 – 376.

Pouscoulous, N. (2011). Metaphor – for adults only? *Belgian Journal of Linguistics*, 25, 64 – 92

Pouscoulous, N. (2014). "The elevator's buttocks": Metaphorical abilities in children. In D. Matthews (Ed.), Pragmatic Development in First Language Acquisition (pp. 239-259). John Benjamins.

Pouscoulous, N. & Tomasello, M. (2011). Very early understanding of metaphors. Poster presented at *the 2011 biennial of the Society for Research in Child Development*, in Montreal, Canada.

Prat, C.S., Mason, R.A., Just, M.A. (2012). An fMRI Investigation of Analogical Mapping in Metaphor Comprehension: The Influence of Context and Individual Cognitive Capacities on Processing Demands. *Journal of Experimental Psychology Learning Memory & Cognition*, 38, 282 – 294.

Pynte, J., Besson, M., Robichon, F.H., Poli, J. (1996). The Time-Course of Metaphor Comprehension: An Event-Related Potential Study. *Brain & Language*, 55, 293 – 316.

Rataj, K. (2014). Surfing the brainwaves of metaphor comprehension. *Poznan Studies in Contemporary Linguistics*, 50, 55 – 73.

Rataj, K., Przekoracka-Krawczyk, A. , Lubbe, R.H.J. (2018). On understanding creative language: the late positive complex and novel metaphor comprehension. *Brain Research*, 1678, 231 – 244.

Regel, S., Gunter, T. C., & Friederici, A. D. (2011). Isn't it ironic? An electrophysiological exploration of figurative language processing. *Journal of Cognitive Neuroscience*, 23, 277 – 93.

Regel, S., Meyer, L., Gunter, T.C. (2014). Distinguishing Neurocognitive Processes Reflected by P600 Effects: Evidence from ERPs and Neural Oscillations. *PLoS one*, 9, e96840.

Reverberi, C., Capitani, E., & Laiacona, M. (2004). Variabili semantico lessicali relative a tutti gli elementi di una categoria semantica: indagine su soggetti normali italiano per la categoria “frutta”. *Giornale Italiano di psicologia*, 3, 497 – 522.

Richards, I.A. (1936). *The Philosophy of Rethoric*. London: Oxford University Press.

Richland, L.E., Morrison, R.G., & Holyoak, K.J., (2006). Children’s development of analogical reasoning: Insights from scene analogy problems. *Journal of Experimental Child Psychology*, 94, 249 – 273.

Rubio Fernández, P. (2007). Suppression in metaphor interpretation: Differences between meaning selection and meaning construction. *Journal of Semantics*, 24, 345 – 371.

Rubio-Fernández, P ., Cummings, C., Tian, Y. (2016). Are single and extended metaphors processed differently? A test of two Relevance-Theoretic accounts. *Journal of Pragmatics*, 94, 15 – 28.

Rubio-Fernández, P. & Grassmann, S. (2016). Metaphors as Second Labels: Difficult for Preschool Children? *Journal of Psycholinguistic Research*, 45, 931 – 944.

Rundblad, G., & Annaz, D. (2010). The atypical development of metaphor and metonymy comprehension in children with autism. *Autism*, 14, 29 – 46.

Ruta, L., Mazzone, D., Mazzone, L., Wheelwright, S., & Baron-Cohen, S. (2012). The Autism-Spectrum Quotient—Italian version: A cross-cultural confirmation of the broader autism phenotype. *Journal of Autism and Developmental Disorders*, 42, 625 – 633.

Rutter, B., Kröger, S., Hill, H., Windmann, S., Hermann, C., & Abraham, A. (2012). Can clouds dance? Part 2: An ERP investigation of passive conceptual expansion. *Brain and Cognition*, 80, 301 – 310.

Sartori, G., Colombo, L., Vallar, G., Rusconi, M. L., & Pinarello, A. (1997). TIB: Test di Intelligenza Breve per la valutazione del quoziente intellettivo attuale e pre-morboso. *La Professione Di Psicologo*, 1, 2 – 24.

Schulze, C., Grassmann, S. & Tomasello, M. (2013). 3-year-old children make relevance inferences in indirect verbal communication. *Child Development*, 84, 2079 – 2093.

Schumacher, P. B. (2011). The hepatitis called ...: Electrophysiological evidence for enriched composition. In J. Meibauer & M. Steinbach (Eds.), *Experimental Pragmatics/Semantics*. Amsterdam: John Benjamins, 199 – 219.

Schumacher, P. B. (2014). Content and context in incremental processing: ‘the ham sandwich’ revisited. *Philosophical Studies*, 168, 151 – 165.

Searle, J. (1979). Metaphor. In A. Ortony (Ed.), *Metaphor and Thought*. Cambridge: Cambridge University Press, 92 – 123.

Simmering, V. R. (2012). The development of visual working memory capacity during early childhood. *Journal of Experimental Child Psychology*, 111, 695 – 707.

Simmering, V.R. & Perone, S. (2012). Working memory capacity as a dynamic process. *Frontiers in Psychology*, 3, 567.

Sperber, D & Wilson, D. (1995) [1986]. *Relevance: Communication and Cognition*, 2nd edn. Oxford: Blackwell.

Sperber, D., Wilson, D. (1996). Fodor's Frame Problem and Relevance Theory (reply to Chiappe & Kukla. *Behavioural and Brain Sciences*, 19, 530 – 532.

Sperber, D., Wilson, D. (2002). Pragmatics, Modularity and Mind-reading. *Mind and Language*, 17, 3 – 23.

Sperber, D. (2005). Modularity and relevance: How can a massively modular mind be flexible and context-sensitive? In Carruthers, P., Laurence, S., Stich, S. (Eds). *The Innate Mind: Structure and Content.* Oxford: Oxford University Press, 53 – 68.

Sperber, D., & Wilson, D. (2008). A Deflationary Account of Metaphors. In R. Gibbs, Jr. (Ed.), *The Cambridge Handbook of Metaphor and Thought*. Cambridge: Cambridge University Press, 84 – 106.

Spotorno, N., Cheylus, A., Van Der Henst, J.-B., & Noveck, I. A. (2013). What's behind a P600? Integration Operations during Irony Processing. *PLoS ONE*, 8, e66839.

Stanovich, K. E., & West, R. F. (1989). Exposure to print and orthographic processing. *Reading Research Quarterly*, 402 – 433.

Steinhauer, K. and Friederici, D.F. (2001). Prosodic Boundaries, Comma Rules, and Brain Responses: The Closure Positive Shift in ERPs as a Universal Marker for Prosodic Phrasing in Listeners and Readers. *Journal of Psycholinguistic Research*, 30, 267 – 295.

Thomason, M.E., Race, e., Burrows, B., Whitfield-Gabrieli, S., Glover, G.H., Gabrieli, JD. (2009). Development of spatial and verbal working memory capacitu in the human brain. *Journal of Cognitive Neuroscience*, 21, 316 – 332.

Tomasello, M. (1999). *The cultural origins of human cognition*. Cambridge: Harvard University Press.

Trick L. & Katz, A. (1989). The domain interaction approach to metaphor processing: Relating individual differences and metaphor characteristics. *Metaphor and Symbolic Activity,* 1, 185 – 213.

Tunteler, E. & Resing, W.C.M. (2002). Spontaneous analogical transfer in 4-year-olds: A microgenetic study. *Journal of Experimental Child Psychology*, 83, 149 – 166.

Tunteler, E. & Resing, W.C.M. (2007). Change in Spontaneous Analogical Transfer in Young Children: A Microgenetic Study. *Infant and Child Development*, 16, 71 – 94.

van den Brink, D., Van Berkum, J.J.A., Bastiaansen, M.C.M., Tesink, M.K., Buitelaar, J.K., Hagoort, P. (2012). Empathy matters: ERP evidence for inter-individual differences in social language processing. *SCAN*, 7, 173 – 183.

Vosniadou, S. (1987). Children and metaphors. *Child Development*, 58, 870 – 885.

Waggoner, J., & Palermo, D. (1989). Betty is a bouncing bubble: Children's comprehension of emotion-descriptive metaphors. *Developmental Psychology*, 25, 152 – 63.

Wagner, S., Winner, E., Cicchetti, D., & Gardner, H. (1981). Metaphorical mapping in human infants. *Child Development*, 52, 728 – 731.

Weiland, H., Bambini, V., & Schumacher, P. (2014). The role of literal meaning in figurative language comprehension: Evidence from masked priming ERP. *Frontiers in Human Neuroscience*, 9, 11 – 12.

Wilson, D., & Carston, R. (2007). A unitary approach to lexical pragmatics: Relevance, inference and ad hoc concepts. In N. Burton-Roberts (Ed.), *Advances in Pragmatics*. Houndmills, Basingstoke: Palgrave, 230 – 260.

Winner, E., Rosenstiel, A. & Gardner, H. (1976). The development of metaphoric understanding. *Developmental Psychology*, 12, 289 – 297.

Winner, E. (1988/1997). *The Point of Words: Children's Understanding of Metaphor and Irony*. Cambridge, MA: Harvard University Press.

Wirth, M., Horn, H., Koenig, T., Stein, M., Federspiel, A. et al. (2007). Sex Differences in Semantic Processing: Event-Related Brain Potentials Distinguish between Lower and Higher Order Semantic Analysis during Word Reading. *Cerebral Cortex*, 17, 1987 – 1997.

Wolff, P., & Gentner, D. (2000). Evidence for role-neutral initial processing of metaphors. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 26, 529 – 541.

Xu, X., kang, C., Guo, T. (2014). The Effects of Gender and Self-Insight on Early Semantic Processing. *PloS ONE*, 9, e114421.

Yang, F.P.G., Bradley, K., Huq, M., Wu, D.L., Krawczyk, D.C. (2013). Contextual effects on conceptual blending in metaphors: An event-related potential study. *Journal of Neurolinguistics*, 26, 312 – 326.

Yang, X., Minai, U., Fiorentino, R. (2018). Context-Sensitivity in the Derivation of Scalar Implicature. *Frontiers in Psychology,* 9, 1720.

Appendix A

Experimental Material used in Study 1.

A.1 - Experimental material used in all trials (metaphoric and literal) of the Metaphor Task in Study 1: a) Target Picture; B) Control Item Target-Literal; C) Control Item Vehicle-Literal.

La macchina con lo zaino/ La macchina con la scatola nel tetto (Eng. The car with the backpack/The car with the box on the roof).

a)

b)

c)

Il gattino coi calzini / Il gattino con le zampe bianche (Eng. The kitten with the socks/The kitten with the white paws).

a)

b)

c)

Le carote con i capelli/ Le carore con le foglie (Eng. The carrots with the hair/ The carrots with the leaves).

a)

b)

c)

La bottiglia col pancione/La bottiglia rotonda (Eng. The bottle with the big belly/The round bottle).

a)

b)

c)

La torre col cappello/La torre col tetto a punta (Eng. The tower with the hat/The tower with the pointy roof).

a)

b)

c)

L'albero con le braccia/L'albero con i rami (Eng. The tree with the arms/The tree with the branches).

a)

b)

c)

Il bicchiere con le antenne/Il bicchiere con le cannucce (Eng. The glass with the antennae/The glass with the straws).

a)

b)

c)

Il cellulare col cappotto/Il cellulare con la custodia (Eng. The mobile phone with the coat/The mobile phone with the cover).

a)

b)

c)

A.2 - Experimental material used in all trials (Same Term and New Term condition) of the Alternative Naming Task.

Micio
(Eng. Kitten)

Stivali
(Eng. Boots)

Spaghetti

Zaino
(Eng. Backpack)

Camion/Macchina
(Eng. Truck/Car)

Libro/Favola
(Eng. Book/Story)

Leccalecca/Caramella
(Eng. Lollipop/Candy)

Colori/Matite
Eng. Colours/Pencils)

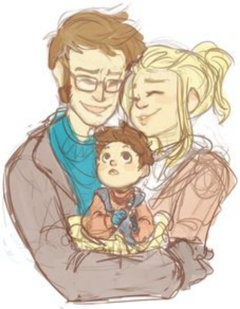
A.3 - Experimental material used in all trials of the Analogy Perception Task. Control 1: *Correct Property/Wrong Object*; Control 2: *Wrong Property/Correct Object.*

Relation: Family

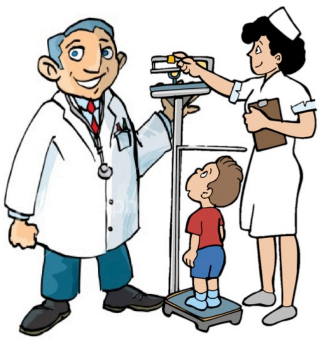
Term 1

Term 2

Target



Control 1



Control 2

Relation: Open

Term 1

Term 2

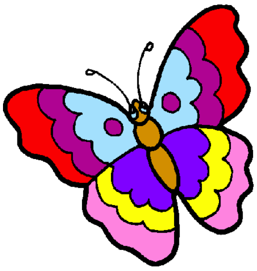
Target

Control 1

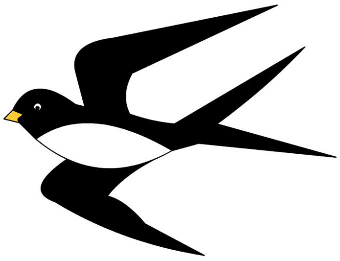
Control 2

Relation: Flying Animals

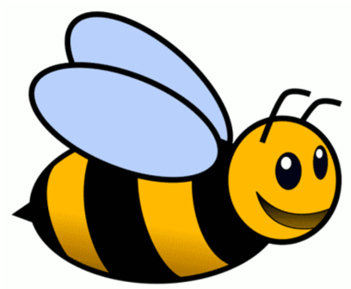
Term 1



Term 2

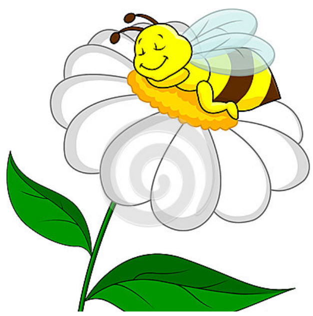


Target



Control 1

Control 2



Relation: Housing

Term 1

Term 2

Target

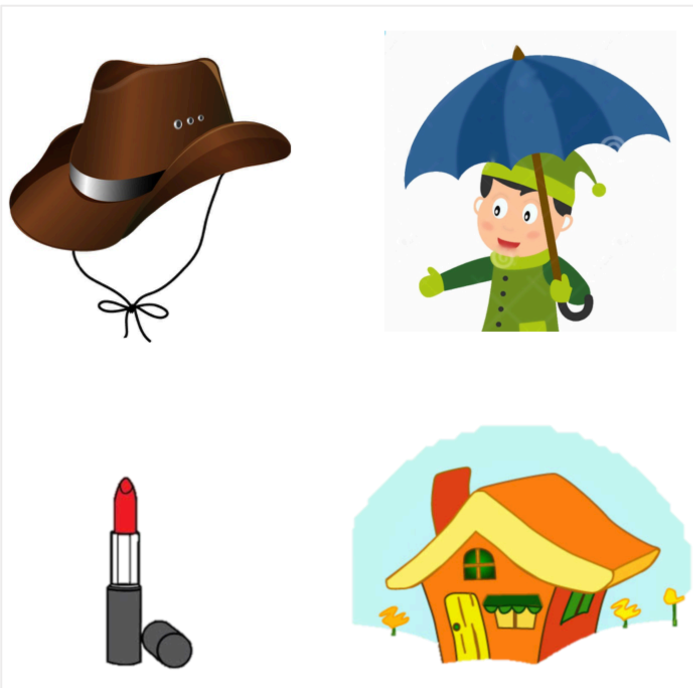
Control 1

Control 2

A.4 – Pictorial material used in the pointing-and-naming picture book that assessed children's comprehension and production of the vocabulary used in the metaphor task. Different pictures for comprehension (top) and production (bottom) were used. The words in bold were assessed.

---

1. The tower with the **hat (cappello)** / The tower with the pointy **roof (tetto)**.



2. The tree with **arms (braccia)** / The tree with the **branches (rami).**

3. The glass with the **antennae (antenne)** / The glass with the **straws (cannucce).**

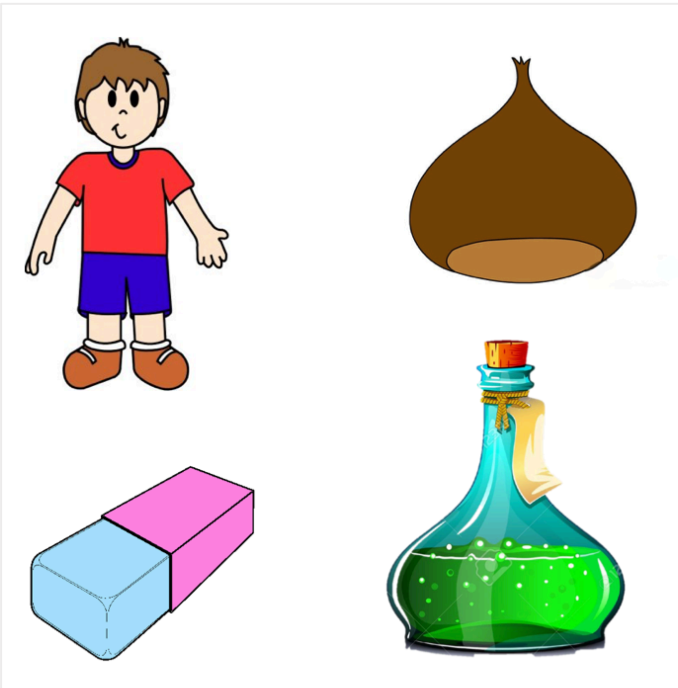
4. The mobile phone with the **coat (cappotto)** / The mobile phone with the **cover (custodia)**.

5. The car with the **backpack (zaino)** / The car with the **box (scatola)** on the top.

6. The kitten with the **socks (calzini) /** The kitten with the white **paws (zampe)**.

7. The carrots with the **hair (capelli)** / The carrots with the **leaves (foglie).**

8. The bottle with the big **belly (pancia)**/ The round **bottle (rotonda)**.



Appendix B

Metaphorical and Literal expressions and material for the selection task used in Study 2 (EEG session).

			SELECTION TASK	
	CONDITION	**TARGET SENTENCE**	**RELATED**	**UNRELATED**
1	met	Dopo una cena a base di aglio gli aliti sono fogne insopportabili.	Puzza	Pelle
2	met	Nei processi importanti certi avvocati sono squali senza scrupoli.	Aggressività	Cartoleria
3	met	Secondo l'opinione del generale quegli eserciti sono dighe insormontabili.	Barriera	Sete
4	met	Quando il tempo stringe gli oratori sono treni ad altà velocità.	Corsa	Elemosina
5	met	Nell'arrampicata ad alta quota gli scalatori sono scoiattoli molto agili.	Scatto	Palazzo
6	met	A volte nell'educazione degli adolescenti gli schiaffi sono medicine efficaci.	Malattia	Stecca
7	met	Specialmente per le coste italiane gli uragani sono carrarmati devastanti.	Distruzione	Università
8	met	Davanti ad una punizione certi bambini sono fontane inarrestabili.	Acqua	Criminalità
9	met	Nel mondo della speculazione finanziaria i banchieri sono vampiri sempre assetati.	Morso	Giraffa
10	met	Nell'utilizzo di rasoi taglienti i barbieri sono chirurghi altamente specializzati.	Precisione	Rivista
11	met	Negli ambienti di lavoro spesso i colleghi sono zecche fastidiose.	Colla	Zampa
12	met	All'occhio di un investigatore alle prime armi certi delitti sono rebus insolubili.	Enigma	Zanzara
13	met	Alle finali olimpioniche di atletica i ginnasti sono grilli agilissimi.	Salto	Laurea
14	met	In certi sport americani i giocatori sono elefanti dalla mole imponente.	Dimensione	Polizia
15	met	Durante gli scandali nazionali i giornalisti sono avvoltoi senza morale.	Fame	Bicicletta
16	met	Nella formazione culturale di ogni persona certi libri sono bussole di grande utilità.	Guida	Musica
17	met	Nell'educazione dei giovani certi maestri sono lanterne che illuminano la via.	Luce	Battaglia
18	met	Nelle campagne elettorali certi politici sono calamite per i voti.	Attrazione	Fotografia
19	met	Nei condomini del centro i portinai sono archivi pronti a spettegolare.	Notizia	Pennarello
20	met	Specialmente per le materie umanistiche certi professori sono enciclopedie zeppe di conoscenze.	Informazione	Palazzo

21	met	Nel combattimento corpo a corpo certi soldati sono leoni senza paura.	Coraggio	Litigio
22	met	Secondo la giornalista di moda quelle acconciature sono cespugli incolti.	Disordine	Ristorante
23	met	Durante la gara di formula uno le automobili sono frecce velocissime.	Movimento	Tabaccheria
24	met	Alla prima dello spettacolo le ballerine sono farfalle molto aggraziate.	Delicatezza	Alimentazione
25	met	Per le ricerche di letteratura certe biblioteche sono miniere ricche di informazioni.	Ricchezza	Automobile
26	met	Per i faccendieri del tribunale le borse sono macigni traboccanti di documenti.	Peso	Cellulare
27	met	Per gli adolescenti amanti del rock certe canzoni sono droghe ipnotiche.	Assuefazione	Valigia
28	met	Per i bambini maltrattati le carezze sono balsami che curano il trauma.	Cura	Stivale
29	met	Nell'ora del rientro dal lavoro le città sono giungle caotiche.	Confusione	Autostrada
30	met	Nelle poesie dei poeti romantici le fanciulle sono rose delicate.	Bellezza	Rivista
31	met	Quando si suda per la tensione le giacche sono forni di calore.	Temperatura	Sciarpa
32	met	Nei bambini appena nati le guance sono pesche liscissime.	Morbidezza	Cesta
33	met	Alle sfilate di alta moda le indossatrici sono bambole di incantevole bellezza.	Perfezione	Irascibilità
34	met	Con certi insegnanti dell'università le lezioni in classe sono sonniferi potentissimi.	Sonno	Sport
35	met	Specialmente nei paesi del terzo mondo certe malattie sono cecchini senza pietà.	Mira	Colluttorio
36	met	Per i distubi del sonno certe melodie sono camomille rilassanti.	Calma	Nave
37	met	Nelle situazioni di crisi economica certe notizie sono terremoti destabilizzanti.	Vibrazione	Bolletta
38	met	Le tasche delle giacche dei nonni sono banche zeppe di monete.	Finanziamento	Dipinto
39	met	Ai matrimoni tradizionali le torte sono montagne di panna e praline.	Grandezza	Alligatore
40	met	Specialmente nel genere femminile certe voci sono trombe squillanti e fastidiose.	Rumore	Ringhiera
41	met	Dopo un anno di duro lavoro i viaggi sono terapie di sicura efficacia.	Benessere	Anfora
42	met	Dopo un grave lutto certi ricordi sono spine dolorose.	Malessere	Vaccino
43	met	Nelle persone creative le menti sono vulcani di idee e progetti.	Energia	Arredamento
44	met	Al giorno d'oggi i passatempi sono oasi dal lavoro e dalla frenesia del quotidiano.	Sosta	Semaforo
45	met	Specialmente nella adolescenza le passioni sono venti passeggeri e instabili.	Moto	Arma
46	met	In un percorso di miglioramento spirituale le virtù sono muscoli da potenziare.	Forza	Cinema
47	met	Anche dopo molti anni i rimproveri dei genitori sono tatuaggi nella memoria.	Eternità	Caldaia
48	met	Per gli animi sensibili le poesie sono sorrisi per il	Felicità	Ascensore

		cuore.		
49	met	I malati di artrosi dicono che le fitte sono coltelli nella carne.	Lama	Orologio
50	met	Specialmente in filosofia certe teorie sono palazzi dalle fragili fondamenta.	Costruzione	Portafoglio
51	met	Specialmente in America certe università sono portaerei che trasportano talenti.	Volo	Matrimonio
52	met	Tutti sanno che le amicizie sono alberi che crescono con il tempo.	Radice	Candela
53	met	Se non curati con farmaci specifici certi malditesta sono trapani continui.	Buco	Lampada
54	met	Nei paesi in via di sviluppo le economie sono bombe ad orologeria.	Scoppio	Poltrona
55	met	Nella memoria dei popoli le stragi sono cicatrici indelebili.	Ferita	Televisione
56	met	Soprattutto nelle culture tradizionali i matrimoni sono querce dalle radici solide.	Solidità	Zaino
57	met	Quando si toccano argomenti delicati le parole sono pietre che pesano sull'animo.	Fastidio	Articolo
58	met	In certi momenti gli occhi sono oceani di profondità infinita.	Vastità	Arroganza
59	met	Dopo un temporale le nuvole sono perle di infiniti riflessi.	Pregio	Allucinazione
60	met	Se negative certe amicizie sono virus che infettano il carattere dei giovani.	Malattia	Piscina
61	met	Nelle patologie psichiatriche le ossessioni sono gabbie che imprigionano.	Sbarra	Uccello
62	met	Nelle scelte di vita i dubbi sono ombre che offuscano la via.	Buio	Amicizia
63	met	Per alcune persone certi profumi sono chiavi per la porta della memoria.	Passaggio	Colazione
64	met	Nei momenti di sconforto i sogni sono ali per viaggiare lontano.	Allontanamento	Biscotto
65	met	Spesso nella vita le delusioni sono zavorre molto pesanti.	Carico	Termosifone
66	met	Negli indecisi le opinioni sono pendoli che vanno di qua e di là.	Oscillazione	Lavoro
67	met	Nel mondo dello spettacolo certe conoscenze sono ascensori per il successo.	Salita	Maleducazione
68	met	Per i soggetti depressi le crisi sono tunnel senza fine.	Oscurità	Eleganza
69	met	Nella vita le paure sono nodi che impediscono di andare avanti.	Freno	Tramonto
70	met	Nei momenti difficili le speranze sono stelle che illuminano l'anima.	Bagliore	Interruttore
71	met	Per molti anziani i figli sono stampelle che sostengono nella vecchiaia.	Appoggio	Mare
72	met	Nelle società multietniche le lingue sono ponti tra le persone.	Contatto	Presunzione
73	met	Nel caso di alcuni autori certi romanzi sono mappe che aiutano ad orientarsi nella vita.	Direzione	Accendino
74	met	Specialmente nelle relazioni sentimentali i torti sono crepe nell'anima.	Rottura	Tegame
75	met	Nel corso della vita le emozioni sono fulmini che sconvolgono la routine.	Apparizione	Cucchiaio

76	met	Quando c'è diffidenza i segreti sono muri tra le persone.	Divisione	Scrivania
77	met	In politica e nel commercio certi discorsi sono fiumi che travolgono.	Corrente	Chiave
78	met	Nella società umana i rapporti sono corde che ci legano agli altri.	Vicinanza	Elicottero
79	met	Senza il perdono i rancori sono tarli che divorano l'anima.	Invadenza	Nutrizione
80	met	Negli animi fragili le tentazioni sono reti molto pericolose.	Imprigionamento	Disorganizzazio ne
1	lit	Illustrando la mappa l'architetto segnala che gli scarichi sono fogne per i rifiuti tossici.	Puzza	Pelle
2	lit	Durante la visita la guida spiega che i pesci nell'acquario sono squali ferocissimi.	Aggressività	Cartoleria
3	lit	Gli ingegneri informano che le costruzioni pianificate sono dighe anti alluvioni.	Barriera	Sete
4	lit	I convogli presenti nel vecchio deposito sono treni in uso cento anni fa.	Corsa	Elemosina
5	lit	Il documentario spiega che i roditori della montagna sono scoiattoli dalla lunga coda.	Scatto	Palazzo
6	lit	Nella cura delle emicranie le sostanze in fase di sperimentazione sono medicine potentissime.	Malattia	Stecca
7	lit	Alla parata militare i blindati più impressionanti sono carrarmati usati in guerra.	Distruzione	Università
8	lit	All'inaugurazione del giardino viene spiegato che i nuovi impianti sono fontane notturne.	Acqua	Criminalità
9	lit	A carnevale i mostri più diffusi sono vampiri dai denti affilati.	Morso	Giraffa
10	lit	I dirigenti sanitari spiegano che quei medici sono chirurghi afferenti all'ospedale.	Precisione	Rivista
11	lit	I parassiti che infestano le piante d'appartamento sono zecche o pidocchi.	Colla	Zampa
12	lit	Nei quiz gli indovinelli misteriosi sono rebus molto difficili.	Enigma	Zanzara
13	lit	I genitori spiegano al bambino che gli insetti del fiume sono grilli agilissimi.	Salto	Laurea
14	lit	Negli zoo asiatici i mammiferi in gabbia sono elefanti provenienti dall'India.	Dimensione	Polizia
15	lit	I predatori che compaiono nei libri di scienze sono avvoltoi della savana.	Fame	Bicicletta
16	lit	Nel kit per il campeggio i dispositivi più costosi sono bussole ad alta precisione.	Guida	Musica
17	lit	Nei velieri vichinghi le lampade più antiche sono lanterne a olio.	Luce	Battaglia
18	lit	Gli oggetti usati per abbellire il frigorifero sono calamite con decori artigianali.	Attrazione	Fotografia
19	lit	Gli schedari presenti negli uffici comunali sono archivi pubblici e consultabili da tutti.	Notizia	Pennarello
20	lit	Nella biblioteca i volumi più antichi sono enciclopedie di inizio novecento.	Informazione	Palazzo
21	lit	Nelle zone equatoriali i felini a rischio di estinzione sono leoni e ghepardi selvatici.	Coraggio	Litigio
22	lit	Al castello abbandonato le siepi di recinzione sono cespugli incolti.	Disordine	Ristorante
23	lit	Per gli indiani d'america le armi più efficaci erano	Movimento	Tabaccheria

		frecce dalla punta affilata.		
24	lit	Nei disegni dei bambini gli animali più colorati sono farfalle multicolore.	Delicatezza	Alimentazione
25	lit	Nel territorio di montagna le cave abbandonate sono miniere di ferro e magnesio.	Ricchezza	Automobile
26	lit	Le pietre da lavorare sono macigni che ricoprivano il lato della montagna.	Peso	Cellulare
27	lit	La polizia spiega che le pasticche colorate sono droghe molto pesanti.	Assuefazione	Valigia
28	lit	Nella cura dei capelli i detergenti più nuovi sono balsami che non danneggiano il fusto.	Cura	Stivale
29	lit	Le foreste dell'America centrale sono giungle selvagge.	Confusione	Autostrada
30	lit	Nei giardini a primavera i fiori più profumati sono rose e gelsomini.	Bellezza	Rivista
31	lit	Nelle cucine americane gli elettrodomestici più richiesti sono forni a micronde.	Temperatura	Sciarpa
32	lit	Ai mercati generali i frutti profumati sono pesche e albicocche di stagione.	Morbidezza	Cesta
33	lit	Tra i regali di Natale i pupazzi più richiesti sono bambole parlanti.	Perfezione	Irascibilità
34	lit	Il farmacista spiega che le pillole prescritte sono sonniferi integrati con antidolorifici.	Sonno	Sport
35	lit	Durante la guerra i militari mandati in avanscoperta erano cecchini di grande precisione.	Mira	Colluttorio
36	lit	Nella bottega dell'erborista le bevande più richieste sono camomille rilassanti.	Calma	Nave
37	lit	L'ente nazionale ha informato che le scosse registrate sono terremoti lievissimi.	Vibrazione	Bolletta
38	lit	Nell'elenco delle società quotate gli istituti ai vertici sono banche con capitali stranieri.	Finanziamento	Dipinto
39	lit	Nelle cartine geografiche rilievi indicati con colori scuri sono montagne di altezza superiore a 2000 metri.	Grandezza	Alligatore
40	lit	Nella nuova orchestra gli strumenti più bizzarri sono trombe bitonali dal suono molto acuto.	Rumore	Ringhiera
41	lit	Nei centri termali i trattamenti a base di fanghi sono terapie molto efficaci.	Benessere	Anfora
42	lit	Nelle piante carnivore le punte delle foglie sono spine acuminate e taglienti.	Malessere	Vaccino
43	lit	Negli atolli tropicali i crateri maggiori sono vulcani ancora attivi.	Energia	Arredamento
44	lit	Durante il safari la guida spiega che le radure in lontananza sono oasi per i nomadi.	Sosta	Semaforo
45	lit	Le brezze che increspano il mare sono venti provenienti da oriente.	Moto	Arma
46	lit	Nel corpo umano le fibre più lunghe sono muscoli striati.	Forza	Cinema
47	lit	Nelle culture aborigene i segni sulla pelle sono tatuaggi di iniziazione.	Eternità	Caldaia
48	lit	Le espressioni più belle sono sorrisi e sguardi sinceri.	Felicità	Ascensore
49	lit	Gli arnesi con manico in legno sono coltelli per conciatori.	Lama	Orologio

50	lit	Nei cantieri di periferia le nuove costruzioni sono palazzi popolari.	Costruzione	Portafoglio
51	lit	Dicono che le imbarcazioni avvistate al largo sono portaerei americane.	Volo	Matrimonio
52	lit	I semi piantati nel giardino sono alberi da frutto e arbusti vari.	Radice	Candela
53	lit	Quegli attrezzi in disuso sono trapani che funzionano a batteria.	Buco	Lampada
54	lit	Nell'attentato alla metropolitana gli esplosivi erano bombe ad orologeria.	Scoppio	Poltrona
55	lit	I segni coperti dal trucco sono cicatrici dovute alle operazioni.	Ferita	Televisione
56	lit	Gli alberi più antichi del parco sono querce millenarie.	Solidità	Zaino
57	lit	I detriti lasciati dal fiume sono pietre e ghiaia di piccole dimensioni.	Fastidio	Articolo
58	lit	In geologia le croste più profonde sono oceani di pietre laviche.	Vastità	Arroganza
59	lit	Le gemme delle collane reali sono perle provenienti dalle Indie.	Pregio	Allucinazione
60	lit	I parassiti delle piante da frutto sono virus molto pericolosi per tutte le coltivazioni.	Malattia	Piscina
61	lit	I contenitori per rettili sono gabbie metalliche molto resistenti.	Sbarra	Uccello
62	lit	Le forme sulle pareti delle grotte sono ombre generate dagli scogli al tramonto.	Buio	Amicizia
63	lit	Nella vecchia ferramenta gli arnesi nascosti sono chiavi da scasso.	Passaggio	Colazione
64	lit	Negli uccelli primitivi le cartilagini esterne erano ali finalizzate al trasporto.	Allontanamento	Biscotto
65	lit	Nelle barche tradizionali i sacchi di sabbia sono zavorre per l'ancoraggio.	Carico	Termosifone
66	lit	Tra i pezzi di artigianato in legno gli orologi sono pendoli a ricarica manuale.	Oscillazione	Lavoro
67	lit	Il capocantiere spiega che i montacarichi sono ascensori per i piani alti.	Salita	Maleducazione
68	lit	Le voragini nella montagna sono tunnel scavati dai soldati della Grande Guerra.	Oscurità	Eleganza
69	lit	Nelle barche certi lacci sono nodi per bloccare le vele.	Freno	Tramonto
70	lit	Quelle luci nel cielo notturno sono stelle di galassie lontane.	Bagliore	Interruttore
71	lit	Quelle assi in legno sono stampelle rudimentali piuttosto scomode.	Appoggio	Mare
72	lit	Le costruzioni in mattoni sono ponti di collegamento tra le sponde.	Contatto	Presunzione
73	lit	Nel museo i documenti più preziosi sono mappe in pergamena.	Direzione	Accendino
74	lit	Quei segni sul muro sono crepe formatesi alcuni anni fa.	Rottura	Tegame
75	lit	Quei bagliori in lontananza sono fulmini che preannunciano tuoni e tempesta.	Apparizione	Cucchiaio
76	lit	Negli alberghi di bassa categoria i divisori tra camere sono muri in calcestruzzo.	Divisione	Scrivania
77	lit	In America del Sud molti canali sono fiumi	Corrente	Chiave

		navigabili.		
78	lit	In magazzino le trecce in canapa sono corde che servono per legare i sacchi.	Vicinanza	Elicottero
79	lit	Quegli insetti nell'angolo del muro sono tarli che aggrediscono legno.	Invadenza	Nutrizione
80	lit	Quei fili sottili sono reti per catturare pesci piccoli.	Imprigionamento	Disorganizzazione