U N I V E R S I T E L U M I E R E – L Y O N 2
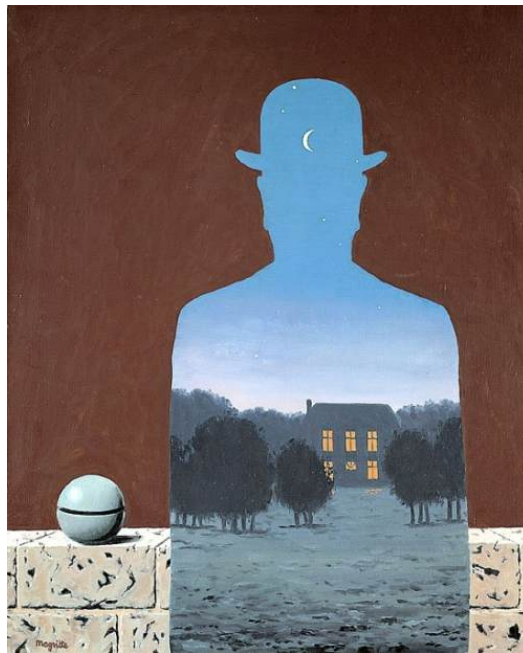
Ecole Doctorale de Neurosciences et Cognition (NSCo)

Laboratoire sur le Langage, le Cerveau et la Cognition, CNRS FRE 3406

AN IRONIC TURN OF EVENTS

Theory of Mind in language processing

Par Nicola SPOTORNO

Thèse de Doctorat de Sciences Cognitives, Mention Psychologie Cognitive

Dirigée par Ira NOVECK

Devant un jury composé de:

Philippe De Brabanter

Zohar Eviatar

Ira Noveck

Jérôme Prado

Deirdre Wilson

René Magritte, L'heureux donateur,

1966, Huile sur toile, 55,5 x 45,5 cm,

Musée d'Ixelles, Bruxelles,

© ADAGP, Paris 2008

Table of Content

INTRODUCTION

I   An idea comes to life in a regular day

One day I went to visit Ira, my supervisor, in Tel Aviv in order to actually write up one of the studies presented here. Hungry, we went off in search of falafel sandwiches, which are typically prepared quickly as customers indicate what sort of vegetables they want to go in their pita. Once inside the store, Ira, whose Hebrew (by his own admission) is fair but in need of improvement, led the way by ordering both his sandwich and beverage. When it was my turn, I followed suit by mostly pointing (while Ira, well, supervised). At one point the server asked us something and our differing reactions proved to be very revealing. While Ira responded by saying "what?" I simply pointed to the lemonade dispenser. That is, while Ira reran the server's question ("what do you want to drink"?) in his mind in order to decipher the linguistic code, the server went on to happily fulfill my order. In this case, I relied on prior evidence and mindreading abilities to surmise the server's question.

As one could see, the situation precluded me from employing my linguistic abilities, leaving me only with my other intention-gathering tools. Yet, despite what would generally be considered a disadvantage, my response was faster than Ira's, an interlocutor with access to both the code and ToM. To put it another way, my ToM operated independently (without relying on the linguistic code) and, in this particular case, this ended up providing me with an advantage for efficiently handling linguistic input in an exchange. As soon as this occurred, both Ira and I recognized that this little story was emblematic of one of the main findings in the research program behind the thesis, which is to show how there are times when features of ToM facilitate or even trump linguistic processing.

Of course, a thesis cannot be reduced to a kind of parable. When the work in this thesis began, it started in fact with several open-ended story lines. One was to determine -- in processing terms -- how ToM interacts with language and, with this in mind, irony was chosen as a topic because it makes manifest (both theoretically and intuitively) that comprehension of an utterance routinely calls for ToM, an idea that has remained submerged in the psycholinguistic literature. At the same time, this choice of topic was risky because this was not an obvious line of research in an extensive psycholinguistic literature on irony (not to mention a companion neurological literature). Instead, one finds a few, largely unsolved debates along with what are ostensibly contradictory data. So, another opening line (of questions) was "can one bring order to this literature"? In the rest of the thesis -- which

includes three scientific articles (of which one is accepted and another submitted) -- I aim to address these multiple story lines as I focus on irony as a means to show how ToM interacts with language processing.

In chapters 1 and 2 we will briefly present some background notions about irony and mindreading abilities (Theory of Mind); in chapter 3, we will confront the mixed results proposed by the psycholinguistic literature, and in chapter 4 we will show, by three reading time experiments, that, taking into account the role that mindreading plays in irony processing, one can explain the mixed results. Chapter 5 will introduce the neuroimaging literature on Theory of Mind and irony, revealing that previous fMRI studies have found very little pointing to Theory of Mind. In chapter 6, we will present a new fMRI study, based on the same paradigm we employed in the behavioral experiments, that reveals a strong engagement of the Theory of Mind network in irony processing. Chapters 7 and 8 will be devoted to the investigation of the time course of irony processing. Time frequencies analysis will reveal interesting insights into the allocation of different cognitive resources *during* the comprehension process of an ironic utterance. In the end, chapter 9 will provide concluding remarks.

1. An ironic turn to Experimental Pragmatics

If one is interested in Experimental Pragmatics and wants to better understand the processes underlying utterance understanding, one ultimately focuses on pragmatic inference making, which is the process that allows one to integrate a linguistic stimulus, i.e. an utterance, with contextual information. This would appear to include a very wide set of phenomena that could include politeness, indirection, honorifics, metaphor, irony, metonymy, humor, scalar implicature, contrastive inference and many others. However, when one looks at the current literature, many of the phenomena investigated experimentally end up concerning particular words in utterances. For example, current research behind in the Experimental Pragmatics framework has been focused on scalar terms (e.g., Bott, Bailey, & Grodner, 2012; Bott & Noveck, 2004; Cummins, Sauerland, & Solt, 2012; Grodner, Klein, Carbary, & Tanenhaus, 2010; I. Noveck, 2001; I. A. Noveck & Sperber, 2004; I. Noveck & Reboul, 2008; Pouscoulous, Noveck, Politzer, & Bastide, 2007), metaphorical vehicles (e.g., Deamer, Pouscoulous, & Breheny, 2010; Glucksberg, 2003; Rubio Fernández, 2007), referential terms (e.g., Davies & Katsos, 2010; B Keysar, Barr, Balin, & Brauner, 2000; Boaz Keysar, Lin, & Barr, 2003; Nadig & Sedivy, 2002; Sedivy, 2003) and presuppositions (e.g., Chambers & Juan, 2008; Chemla & Schlenker, 2012; Moxey, 2006). One can detect that Experimental Pragmatics has tended to become microscopic. This tendency is not entirely surprising given that throughout its history, pragmatics has struggled to show that syntactic and semantic aspects of linguistics are not "the only games in town." If one wants to show, say, that truth-conditions are not the only purpose of linguistic communication (e.g., Austin, 1962; P. Grice, 1975; P. H. Grice, 1989; Searle, 1969, 1979; Sperber & Wilson, 1986), it is justified for experimental research to be focused on the lexical level because one can see how even single words are "context dependent."

However, this emphasis on the word-level comes with drawbacks because a pragmatist would insist that her unit of analysis is the utterance and not individual subunits within the utterance. After all, Pragmatics is the study of language in real context of use and utterances, more than individual words, are generally considered to be the units of daily communication. This observation leads to the following question: Are there phenomena, that have concerned psycholinguists, that can receive sentence-wide treatment?

One phenomenon that -- prima facie -- requires a listener to understand an utterance in its entirety is irony, which brings with it a whole host of features that makes it ideal for someone interested in Experimental Pragmatics. Let me point out two. First, it shows how a given utterance can have a thoroughly different meaning as a function of context. To see that, let me present an example that will be using throughout this thesis. Imagine an opera singer who says (1) to her colleagues after being part of what was clearly an awful performance.

1 "Tonight we gave a superb performance!"

Obviously she is being ironic and of course, if the performance was astounding instead of awful, the exact same sentence (1) is no longer ironic; it is just a literal remark on the state of affairs. To appreciate how irony is generally unique in readily providing two interpretations compare it to the case of metaphor. Consider a remark such as "Mary is a bulldozer." It is very difficult to see how the exact same sentence that refers to the same entity (i.e., "Mary" who is a person and "bulldozer" which refers to a machine) can have a meaningful interpretation other than a metaphorical one (for an extensive analysis of metaphor see for example Carston, 2010; Sperber & Wilson, 2008; Wilson & Carston, 2006). So one advantage that irony has is that it is always open, as far as we can tell, to at least two interpretations (as a function of context).

A second feature of irony is that it clearly shows that the linguistic code underdetermines the intended meaning of an ironic utterance; i.e., lexical decoding can go only so far in utterance understanding. Examples like the one in (1) show that the literal reading is false in context even though the listener would understand it. This underlines how the linguistic stimulus can be seen as just a cue aimed to clarify the speaker's intention. In order to fill in the gap between the code and the complete meaning of the utterance one has to critically interpret the speaker's intentions because an ironic utterance is designed to convey the speaker's attitude about a state of affairs.

These two theoretical features correspond with two experimental advantages. One is that a target sentence, e.g., the utterance in (1) can be both an experimental object as well as a control and, experimentally speaking, this situation is ideal. All one needs is a small change in a context to view an ironic sentence as a literal one (and vice versa).

The other is that irony gives one the chance to investigate what cognitive resources are engaged, other than the language "faculty," during the comprehension of an utterance and, when it comes to irony (indeed for pragmatics more generally), one can see how access to a

speaker's intention is vital. From a psychological point of view, this mindreading ability, which usually falls under the category of Theory of Mind (ToM), is now a well developed area. So, while ToM activity is involved in every communicative act (making its experimental isolation difficult), the comparison of irony to (sincere) literal readings makes for a great test bed because an irony's intended meaning arguably maximizes a listener's need to mindread in language comprehension.

2. Irony, attitude and Theory of Mind

If an experimental pragmatist were to investigate irony from scratch today (without knowing anything about current empirical findings) she would a) look at the given pragmatic theories on irony; and; b) try to determine how the ToM literature -- which has grown exponentially and become more sophisticated over the last two decades -- fits into her investigations. Only then would one determine how well current data correspond with theoretical accounts. When one does this one notices that the results in the psycholinguistic studies largely do not illuminate questions on ToM. This is true with respect to the neuroimaging research as well. To appreciate just how much disconnect there is between theory and data with respect to irony, below I present (briefly) some basic background about theories of irony and ToM.

2.1   Irony in theories of Pragmatics: The role of attitude ascription

From classical antiquity to Gricean pragmatics (e.g., from Quintilian, first century A.D./1921; to Grice, 1989), a rich literature has developed in linguistics, rhetoric and literary studies on the nature and uses of irony. This literature a) accepts as fundamental the basic assertion of classic rhetoric that misdirection is the crucial feature of irony and; b) labels ironic meaning as "the opposite of what the speaker said." This label has become the traditional definition of irony, i.e. the one found in dictionaries. However, if one looks carefully at the example of the two singers one can see that this description of irony is not satisfying and for two reasons.

One is that if the singer just wanted to communicate to her colleague that their performance was awful she could have said so. Indeed, from the point of ostensive-inferential communication (e.g., Carston, 2002; P. Grice, 1975; P. H. Grice, 1989; Sperber & Wilson, 1986; Wilson & Sperber, 2004), it is generally assumed that the speaker would do her best to convey the intended meaning and so she should avoid effortful/deviant formulations if they are not necessary. So, this definition for irony does not come with any obvious added value. Second, the singer's colleague probably knows that the performance was bad (or she would
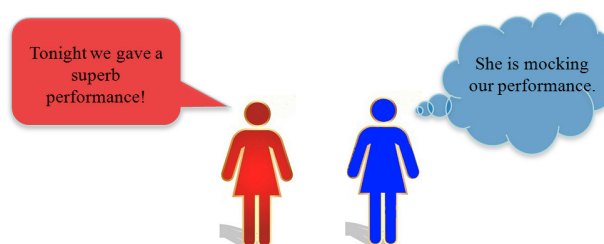
not be able to get the ironic meaning) so simply stating the obvious (albeit through some figurative technique) also does not add anything new. In the end, the definition of the ironic meaning as just "the opposite of what the speaker said" turns irony into an effortful and meaningless communicative act.

What is missing in the above description? An ironic interpretation critically involves the ability to ascribe attitudes. In fact, the main theories of Pragmatics assert that a speaker who uses an ironic remark does so in order to convey her personal attitude.

2.1.1 Post-Gricean approaches: The Echoic theory of irony

Paul Grice (1967/89), who analyzed irony as an instance of figurative language while not distinguishing the comprehension processes of irony from, say, metaphor and who is one of the most eminent exponents of the classic account of irony, still did point out that irony involves a "hostile or derogatory judgment or a feeling such as indignation or contempt" (Grice, p. 53).

Sperber and Wilson, who challenged the classic account, went further with their Echoic theory of irony (e.g., Sperber & Wilson, 1981; Wilson, 2006; Wilson, 2009; Wilson & Sperber, 2012a). They proposed that verbal irony is a subset of an *attributive* use of language. One of the most common sorts of attributive use of language is the ascription of a thought to someone else, as in: "John thinks that it's Monday" This sentence is not directly about the actual day, but it is about another thought, which resembles in content, that the speaker attributes to some source other than herself, i.e. John. The *echoic* use of language is a subcategory of the attributive use. In this case the main message the speaker wants to convey is not the content of the attributed thought but her own attitude or reaction to it. To make a long story short, consider the following exchange, with two identical responses to (2):

(2)   John says: "It's Monday"

(3)   Mary says: "John thinks that it's Monday." (Mary is informing someone else about John's thought because John is the only one with a calendar)

(4)   Mary says: "John thinks that it's Monday!" (actually, is Tuesday but John is so drunk to have forgotten what day it is)

(2) is an ordinary *descriptive* use of language with which John describes a state of affairs. (3) is an *attributive* use of language because Mary's utterance does not refer directly to the state of affairs but to John's thought. (4) is an *echoic* use of language because the content of Mary's sentence is not so important, it is just used to convey Mary's reaction to John's utterance. Verbal irony is a subtype of echoic use in which the utterance conveys a skeptical, mocking attitude about a thought that is attributed to someone else. To make this even clearer, consider an example from Sperber and Wilson (1981) in which someone says (5) to a friend who is fanatical about keeping her flowers watered (so the listener is someone who waters her flowers with great regularity):

(5)   "Did you remember to water the flowers?"

Sperber and Wilson argue that:

"The speaker (of 5) is not interested in the answer; he is much more likely to have asked the question precisely to highlight its irrelevance and the pointlessness of asking or answering it in the circumstances. If we also suppose that the hearer is fanatical about keeping the flowers watered, (5) will have the further implication that the question is USUALLY pointless, and that the hearer's obsession is ridiculous. Thus, what the speaker actually communicates is not the question (5) itself, but an attitude to it and to the state of mind that might give rise to it." (Sperber & Wilson, 1981, 302)

2.1.2 Post-Gricean approaches: The pretense account of irony

In response to Jorgensen et al. (1984), which experimentally supports Sperber and Wilson's account, Herbert Clark and Richard Gerrig, proposed the Pretense Theory (Clark & Gerrig, 1984), which contrasts with the Echoic theory of irony. Clark and Gerrig claimed that the basic mechanism behind irony is not the attributive use of language but a speaker's fake communicative act. The main idea is that the speaker of an ironical utterance is not herself performing a speech act but pretending to perform one, in order to convey a mocking, skeptical or contemptuous attitude to the speech act itself. Like an actor on a stage, the ironic

speaker stops to be herself for a moment and becomes a character who asserts something that is clearly false or inappropriate. The fictive scenario should allow the speaker to take distance from the content of the utterance said in order to reveal the speaker's mocking attitude about it. Clearly the "mise en scène" works only if the audience is able to understand the pretense.

> *"Ironists can pretend to use the words of any person or type of person they wish, just as long as they can get the intended audience to recognize the pretense and, thereby, their attitude toward the speaker, audience, and sentiment of that pretense."* (Clark & Gerrig, 1984, 124)

While the Echoic account remains one of the most plausible explanations of irony (Wilson 2009, Wilson and Sperber, 2012) the "pure" Pretense theory has given way to hybrid accounts that try to fit attributive elements of the Echoic theory into a pretense-based account (see for example, (e.g., Currie, 2006; Kumon-Nakamura, Glucksberg, & Brown, 1995; Recanati, 2007; Walton, 1990). While the thesis will not pursue the differences among the theories (which remain critical), it is important to point out that none of these accounts has dropped the idea that the communication of the speaker's attitude is the hallmark of irony. All the protagonists in the debate take for granted the notion that the message conveyed by an ironic remark is the speaker's mocking, skeptical or contemptuous attitude.

2.2 Theory of Mind

In Premack and Woodruff's seminal paper "Does the chimpanzee have a 'theory of mind'?" (1978), they described how Theory of Mind (ToM) refers to the ability to impute mental states to oneself and to others. The ability to make inferences about what others believe is clearly a crucial component of humans' social skills because it allows one to explain and predict others' behavior. To what extent we share this ability with other animals is controversial, but it is generally accepted that ToM is an evolved psychological ability, which is most highly developed in humans. Our mindreading abilities are specialized for the rapid attribution of beliefs, intentions, desires, or knowledge to others and ourselves and in the spontaneous understanding that others have mental states that may differ from our own.

Mindreading pervades human cognition beyond the border of communicative exchanges with one another. When watching others, we automatically interpret their current

behavior in light of the intentions they might have. If we see a man in a queue at a coffee machine, for example, we will be inclined to explain his behavior by the fact that he wants a coffee, and this attribution of intention will lead us to predict that he will put coins into the machine and that he will choose his favorite coffee. Attributing mental states and predicting behavior on that basis can thus be seen as our most natural way of grasping the social world.

2.2.1 The intentional stance

One developed theoretical idea behind the notion of Theory of Mind is Daniel Dennett's "Intentional stance" (e.g., Dennett, 1987, 2009). In brief, the intentional stance is a strategy that we use when we try to explain the behavior of something on the bases of beliefs and desires. Dennett defined the notion of intentional stance in opposition with two other kinds of stances, namely, the physical stance and the design stance. In Dennett's words:

> "[the physical stance] *is simply the standard laborious method of the physical sciences, in which we use whatever we know about the laws of physics and the physical constitution of the things in question to devise our prediction.."* (Dennett, 2009, p. 2)

On the contrary, in order to predict the behavior of an artifact one often has to apply a design stance:

> *"Alarm clocks, being designed objects (unlike the stone), are also amenable to a fancier style of prediction—prediction from the design stance. Suppose I categorize a novel object as an alarm clock: I can quickly reason that if I depress a few buttons just so, then some hours later the alarm clock will make a loud noise. I don't need to work out the specific physical laws that explain this marvelous regularity; I simply assume that it has a particular design—the design we call an alarm clock—and that it will function properly, as designed."* (Dennett, 2009, p. 2-3)

The intentional stance is a subspecies of the design stance that we employ to explain the behavior of a rational agent:

"(...) a subspecies of the design stance, in which the designed thing is treated as an agent of sorts, with beliefs and desires and enough rationality to do what it ought to do given those beliefs and desires." (Dennett, 2009, p. 3)

The category of "what can be considered a rational agent" is so broad that it can be divided in two sub categories, one containing first-order intentional systems, and the other containing second-order intentional systems.

"A first-order intentional system is one whose behavior is predictable by attributing (simple) beliefs and desires to it. A second-order intentional system is predictable only if it is attributed beliefs about beliefs, or beliefs about desires, or desires about beliefs, and so forth." (Dennett, 2009, p. 10)

However, the attribution of intentions to a piece of furniture of the world can just be a useful heuristic for understanding its behavior. Dennett makes this point with the example of a chess-computer-game (Dennett, 2009). One can easily explain how the computer *chooses* the most *rational* move and how it reacts and *reflects* on our moves, but, in the end, the behavior of the computer is at the same level of behavior as the alarm clock. Nevertheless, an explanation of the computer's moves just in terms of design-stance would be both less intuitive and less efficient than the intentional interpretation.

It follows that the intentional stance can be considered a neutral notion that can help one to explain some kinds of behaviors. Several researchers, including David Premack,

Figure 2. Examples of the three stances proposed by Dennett.

have tried to determine the necessary and sufficient conditions for a true/complete intentional system. Premack (1983) proposed, for example, that just second-level intentional systems,
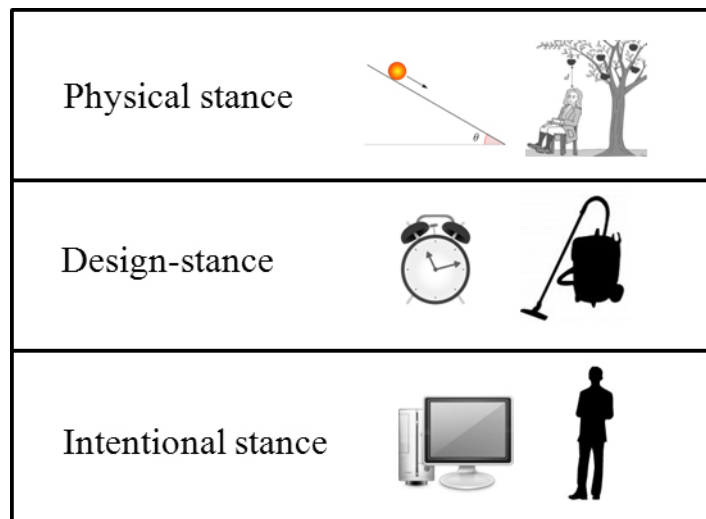
which are systems that are able to attribute to themselves and to others beliefs about beliefs, are the only ones that can really be considered intentional agents.

2.2.2   The golden test

We are clearly moving toward a definition of ToM, but how can one test whether an agent is able to attribute beliefs about beliefs to herself and to others? The most prototypical test that has been designed to discriminate between those having ToM abilities and those who do not is the *false belief task*: passing this test has been taken to mean that one is able to represent others' mental states and failing the test has been taken as evidence that one is impaired in representing others' mental states. In false belief tasks (FBTs), a participant is typically put in a position where she needs to take someone's belief into account so as to provide the appropriate answer to a question. Crucially, the beliefs she needs to take into account differ from her own, and they are also in contrast with a current state of affairs (hence the term "false" belief). Wimmer and Perner (1983) developed the first version of the false belief task, which falls under the category of "unexpected location tasks" because the location of an object is changed during the absence of a character who then has to look for the object when she returns. In this way, the paradigm creates a mismatch between the beliefs of the subject and the beliefs of the character about the location of the hidden object. The original version has been followed by a rich series of variants, the most common of which is the 'Sally-Anne' paradigm. Children are told a story involving two dolls, Sally and Anne, who play with a marble. Sally puts the marble away in a basket and leaves the room. In Sally's absence, Anne takes the marble out and plays with it. Once she has finished playing, she puts the marble away in a box. Sally returns and the child is asked where Sally will look for the marble. The child passes the task if she answers that Sally will look where she first put the marble; the child fails the task if she answers that Sally will look in the box where the marble indeed is (for descriptions of other similar tasks, see Baron-Cohen, 1995).
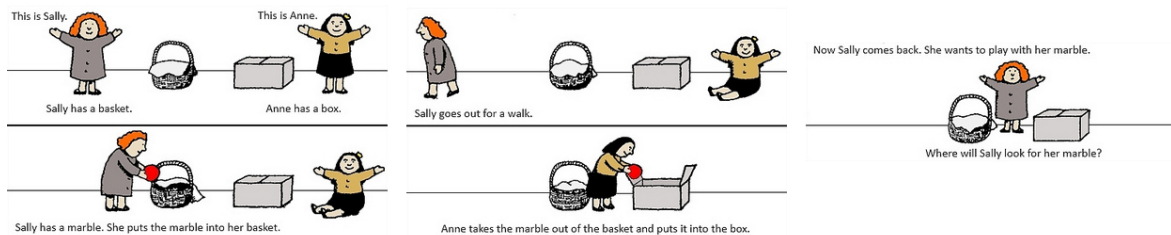
Figure 3. Example of FBT in one of its most common version: the Sally-Anne Task.

Following experimental results from tasks of this type, the following three conclusions have become standard in experimental psychology: First, ToM is specifically human and does not appear to be part of other animals' psychological abilities (including non-human primates). Second, typically developing children master ToM reasoning from around the age of four and rarely succeed in false belief tasks before that age (for a review, see Wellman, Cross, & Watson, 2001). Third, individuals with autism spectrum disorders (ASD) face important difficulties with false belief tasks, which can be taken as evidence of a specific ToM impairment in the condition (Baron-Cohen, 1995).

2.2.3 Challenges to the false belief test

Over the past decade, however, each of these conclusions has been challenged and current knowledge about ToM in primates, typical development and atypical development has been reviewed considerably as a result. First, researchers on primate cognition have accumulated a number of findings suggesting that apes are not fully blind to others' mental states. In experimental settings, there is now robust evidence showing that chimpanzees are able to grasp actions, underlying goals and intentions and that they understand that others perceive the world in a way that may differ from their own (for a review of 10 relevant studies see Call & Tomasello, 2008).

Second, several studies on infants and children younger than the critical threshold of four years old have revealed that they, too, can represent other people's mental states before they pass the FBT (for a review see Baillargeon, Scott, & He, 2010; but see Ruffman & Perner, 2005 for a different prospective). For example, infants who are 15 months old (Onishi & Baillargeon, 2005) or even 13 months old (Surian, Caldi, & Sperber, 2007) are surprised when the actor's behavior does not match their true or false belief regarding the situation. In

addition, recent research in developmental pragmatics has shown that preverbal infants spontaneously take their audience's perspective into account (e.g., Bloom, 2000; Buttelmann, Carpenter, & Tomasello, 2009; Nurmsoo & Bloom, 2008; Southgate, Chevallier, & Csibra, 2010). Moreover, the false belief paradigm engages several abilities other than ToM and so failures in passing it can be due to an immature ToM or to other factors such as verbal demands, executive function demands, expertise or attention levels.

Third, assuming a strict link between Autism Spectrum Disorders and failure in the false beliefs tests has also been questioned. Several studies challenge, not only the validity of the FBT as golden test for ASD but also, the universality of the ToM account of ASD (e.g., Frith & Happé, 1994; Happé, Ronald, & Plomin, 2006; Ozonoff, Pennington, & Rogers, 1991; Pennington, 2006; Chevallier, Noveck, Happé, & Wilson, 2011;). In addition, a recent account called "The social motivation theory of autism" (see for example, Chevallier, 2012; Chevallier, Kohls, Troiani, Brodkin, & Schultz, 2012) proposes that behind deficits in cognitive abilities such as ToM and executive functions are motivational factors that play a major role in ASD. In fact, diminished social orienting, social reward and social maintaining are found in autism and can account for several characteristics of the disorder including difficulties in the development of social cognitive skills (see Chevallier et al., 2012 for a review).

Regardless of the limits of the False Belief Task, the experimental research on ToM has shown to what extent the ability to input to oneself and to the others mental states allows us to explain and predict the behavior of intentional agents. Theory of Mind is clearly a building block of social cognition and the investigation of the interaction between mindreading and linguistic processes can allow us to shed light on the critical human ability that allows us to interact with one another.

IRONY IN PSYCHOLINGUISTICS

3.   The debate around irony processing

When one turns to the psycholinguistic literature one can see that there is a sort of disconnect between the theoretical investigation (e.g., the Echoic account or the pretense-based accounts) and the experimental research. One can find three main lines of research in the literature. I. There is a tradition of off-line studies that investigate a set of contextual variables (e.g., speaker's identity or the relationship between interlocutors) affect the comprehension of an ironic remark (e.g., Katz & Lee, 1993; Katz & Pexman, 1997; Pexman & Olineck, 2002; Pexman, Whalen, & Green, 2010). II. A few papers between the '80's and the beginning of the '90's (e.g., Gibbs, 1986; F. G. Happé, 1993; Jorgensen, Miller, & Sperber, 1984; Keenan & Quigley, 1999) tried to test predictions of theories from Pragmatics. III. The great majority of the on-line studies try to determine whether or not the literal meaning of an ironic utterance is processed before its figurative meaning.

3.1   What makes an utterance more or less ironic

The circumstances in which one makes an ironic remark strongly affect the listener's comprehension and appreciation of it. For example, if John's friend Mary consistently inserts an irony into her discourse, John will very likely recognize almost every one of her ironic remarks. In contrast, if a friend very rarely uses irony, it would probably be hard to detect. In the same way a sentence like: "You are so punctual!" can be immediately interpreted as ironic if the person is late by an hour,while it would be hard to grasp the irony if she is just five minutes late.

Several studies have systematically investigated the circumstances surrounding irony understanding (e.g., Colston, 2002; Gerrig & Goldvarg, 2000; Ivanko & Pexman, 2003; Katz & Lee, 1993; Katz & Pexman, 1997; Kreuz, Kassler, Coppenrath, & McLain Allen, 1999; Kreuz & Link, 2002; Pexman & Zvaigzne, 2004; Pexman & Olineck, 2002; P. Pexman et al., 2010; Roberts & Kreuz, 1994; Slugoski & Turnbull, 1988; Toplak & Katz, 2000). Pexman and Olineck (2002), for example, investigated whether a speaker's occupation is a cue to ironic intent when statements such as "You are a wonderful friend," are potentially ironic insults and sentence like "You are a terrible friend" are potentially ironic compliments. They present subjects stories like (2a) and (2b):

2a Laura and her friend, *a scientist / cab driver* (italic added), were having coffee at a local café. They were talking about a date that Laura had been on the night before. Laura and her date had gone for dinner and dancing at a new club downtown. The *scientist / cab driver* said: "That sounds pretty exciting."

2b Laura and her friend, *a scientist / cab driver* (italic added), were having coffee at a local café. They were talking about a date that Laura had had the night before. They had watched old reruns of cartoons all evening. The *scientist / cab driver* said: "That sounds pretty exciting."

The authors selected, through a pretest, occupations that are considered to be more "sarcastic," e.g., comedians and cab drivers, and occupations that are considered to be "non-sarcastic" like army sergeants and scientists. Results (from two experiments) revealed that target sentences are evaluated as more sarcastic when they are said by a "sarcastic speaker" than by a "non-sarcastic" one and that "sarcastic speakers" are thought to convey a mocking attitude to greater extent than "non-sarcastic speakers." Results from a third experiment showed that stereotypes about a speaker's occupation involve particular types of information in the context of potentially ironic speech: a speaker's perceived tendencies to be humorous, to criticize, to be sincere, as well a speaker's perceived education level.

Studies like Pexman and colleagues' revealed the extent to which contextual information (considering a broad definition of context that includes not only the linguistic context but also background knowledge, such as stereotypes) affects the comprehension and appreciation of an ironic remark. Moreover, the results of these studies show that researchers, when they design their experiments, have to take into account features such as the speaker's identity and the relationship between the interlocutors in order to avoid having these features affect the results.

This line of research does not systematically make reference to theories of Pragmatics[1], but their work pivots around the notion of attitude ascription. Considering that their rating studies ask, for example, to what extent the speaker is mocking the interlocutor

---

[1] Pexman & Olineck (2002) compared their results with the theoretical positions (e.g., the Echoic account and pretence-based proposals) but this is not the focus of their paper and they found just partial overlaps with some of them (e.g., (Colston, 2001; Utsumi, 2000)

(e.g., Pexman & Olineck, 2002), one can say that they take for granted that attitude ascription plays a critical role in irony processing. However, none of these studies directly investigate what happens during the online processing of an ironic utterance.

3.2   The hallmark of irony

Just after Sperber and Wilson took some distance from the classic account with their Echoic mention theory[2] (Sperber and Wilson, 1981; Sperber and Wilson, 1986), a couple of studies have tried to test predictions based on their account (e.g., Jorgensen et al., 1984; Keenan & Quigley, 1999; Happé, 1993; Gibbs, 1986). Most notably, Jorgensen, Miller & Sperber (1984) led the way. They created six frameworks from which they created two versions, such as in (3a) and (3b).

3a   The party was at the Clarks', but Joe didn't know where Mr. Clark lived. "It's on Lee Street," Irma told him. "You can't miss it." But Joe did miss it. He never would have found it if Ken hadn't seen him wandering down the street and led him to the Clarks' apartment. They lived over a store, and their apartment door was right on the sidewalk. Irma was already there when they arrived. "You're late," she called to Joe.
"The Clarks have a beautiful lawn," he replied.

3a   The party was at the Clarks', but Joe didn't know where Mr. Clark lived. "It's on Lee Street," Irma told him. "*It's the house with the big maple tree on the front lawn.* (italic added) You can't miss it." But Joe did miss it. He never would have found it if Ken hadn't seen him wandering down the street and led him to the Clarks' apartment. They lived over a store, and their apartment door was right on the sidewalk. Irma was already there when they arrived. "You're late," she called to Joe.
"The Clarks have a beautiful lawn," he replied.

---

[2] As outlined in Wilson & Sperber (2012) in those early days of relevance theory, they used 'mention' (in an extended sense of the term) to describe what we would later call 'interpretive use'. In the first edition of Relevance (1986), they stop to use 'mention' and have talked since then of the 'echoic' theory of irony.

They argued that if the classic account were correct the sufficient condition for an irony is that "the speaker manifestly expects the hearer to realize that the speaker believes the opposite of that proposition" (Jorgensen et al. 1984, 116). And both (3a) and (3b) respect this condition. In contrast the Echoic account requires that the following condition be fulfilled:

> The prepositional content of the utterance literally understood matches at least in part that of some identifiable utterances, thought, intention, expectation, or norm which it can be taken to echo (page 116).

Only the stories like (3b), in which the antecedent is provided in Irma's first utterance ("*It's the house with the big maple tree on the front lawn.*"), fulfill both criteria. Two questions were asked after each story: one was to ensure that the subject had read the story (e.g., "Where was the party held?"); the other was intended to elicit judgments of irony or sarcasm (e.g., "Why did Joe say, 'The Clarks have a beautiful lawn'"?). The Echoic theory predicts that subjects should easily interpret "The Clarks have a beautiful lawn" as ironic in (3b) while participants may find the same sentence harder to understand in (3a). The results of the study confirmed their prediction, providing the first set of data in favor of the Echoic theory of irony.

In 1993 Francesca Happé wrote a paper that has, rightly, had a great impact on both theory of communication and theories on autism. She argued that Autism can provide a good testing ground to investigate the role that mindreading plays in language processing because, as we have already underlined in the chapter 2, people with ASD manifest severe impairment in Theory of Mind abilities and also difficulties in communication. Happé recognized that Relevance Theory (e.g., Sperber and Wilson, 1986/1995, Sperber and Wilson, 2004; Wilson & Sperber, 2012) makes explicit a role for comprehension of intentions in human communication (see also, Sperber & Wilson, 2002) and that it thus provides a theoretical framework to test and compare the deficits in mindreading and communicative abilities that had been already reported in studies on ASD.

From Relevance Theory, Happé extracted a scale of linguistic phenomena that should correspond to the different levels of mindreading abilities.

> 0th. *Simile* should be understood at a purely literal level. "He was like a lion" should not differ from "He was like his father" because in both cases the hearer has just to decide in what respect there is a similarity. Therefore, even autistic

people who completely lack a Theory of Mind (i.e., people who do not pass the first level version of the False Beliefs Task) should be capable of using and understanding similes, since a literal interpretation will suffice.

1st. *Metaphor* should require some understanding of intentions. In a metaphor the propositional form of the utterance is a more or less loose interpretation of the speaker's thought (see, for example, Carston and Wilson, 2007 and Sperber and Wilson 2006 for the relationship between loose talk and metaphor). Therefore metaphors should require a first-order Theory of Mind to be properly understood. Metaphor has elements of misdirection in the way that False Beliefs Test has; in both cases the speaker / actor's mental states are crucial, and the "ground level", which is the reality in the FBT and the literal meaning of metaphor, is not sufficient for understanding the situation.

2nd. *Irony* should be more demanding because the hearer has to get a thought about an attributed thought, engaging second-order Theory of Mind. In this case the parallel is with the resolution of more complex (second-order) versions of the FBT in which, for example, Anne changes the location of the marble thinking that Sally cannot see, but Sally, looking from the peephole of the door, is observing the action. When Sally comes to the scene, the subject is asked where Anne thinks that Sally will look for the marble. So the subject has to be able to reflect on the beliefs of a character about the beliefs of another character.

The results of Happé's study verified her predictions revealing that a) ASD people who fail the standard FBT are not able to understand both metaphor and irony, but perform well with similes, b) ASD people who pass first-order FBT's but not second-order FBT master the comprehension of metaphors but not of irony, and that; c) only ASD people who pass second-level FBT's are able to properly understand ironic remarks. In addition a further experiment, which was included in the same paper (Happé, 1993), supported the same predictions showing that only typically developing children who pass the second-order FBT are able to correctly understand ironies.

Happé's study provides strong evidence in favor of Sperber and Wilson's analysis of linguistic communication, which includes the Echoic account of irony under the umbrella of Relevance theory (but see section 2.2.3 for challenges to ToM accounts of Autism with respect to FBT's). In any case, this study, as well as Jorgensen et al.'s (1984), cannot provide

insights into the actual comprehension process of irony because it is based on off-line experiments.

In 1986, Raymond Gibbs conducted a reading time experiment that, as far as we know, is the only on-line study on irony that tries to test predictions from the Echoic theory on irony processing. Gibbs had ironic statements serve as concluding remarks in two sorts of (slightly different) contexts that he manipulated following Jorgensen et al.'s (1984) strategy. In one context, there was an explicit mention of a feature that would serve as the basis of an ironic remark. In the Non-echoic version of the story, the sentences containing the explicit mention of the advertisement were excised. Gibbs found reading time reductions on the ironic-target item when there had been an explicit mention (e.g. of a mentioned advertisement) earlier. In other words, in comparison to cases that had no prior explicit mention of a thought that was later echoed in the character's remark, reading time was faster when there *was* an explicit basis to the ironic remark earlier in the story. The data from this experiment served as the basis for claiming that echoic-mention facilitates irony processing. Unfortunately, Gibbs's paradigm suffers from several methodological issues that have been already partially noticed by Rachel Giora (1995) and that will be extensively described in the next chapter (see the paper: "Shouldn't irony be effortful?" in Chapter 4). To sum up, the evidence provided by Gibbs (1986) in favor of the Echoic theory a) supplements Jorgensen et al.'s and b) are arguably not entirely reliable.

3.3   The debate about the literal meaning of an ironic remark

Given standard interpretations of Grice's model, which assumes that listeners detect a violation of a maxim in order to produce implicatures, researchers have been determined to find out whether reading an ironic remark (as measured by, say, its speed) requires more effort than its literal control. This had led to some controversy. On the one hand, there are data showing that irony appears to require more effort to be understood than its literal counterpart. For example, reading time and judgment time studies (e.g., Dews & Winner, 1999; Schwoebel, Dews, Winner, & Srinivas, 2000) showed that a sentence requires more effort to be processed when it was meant ironically than when it was meant literally. Similarly, Filik and Moxey (2010) reported longer latencies for ironic readings over literal readings of the same target sentences in an eye-movement study. On the other hand, there is also evidence indicating that ironic utterances are integrated just as easily as their literal

counterparts (e.g., Gibbs, 1986; Ivanko & Pexman, 2003; Katz, Blasko, & Kazmerski, 2004). Gibbs's work was especially noteworthy for it heralded a Direct Access account, which claimed that Grice's Standard Pragmatic Model (SPM) must be wrong on account of participants' equal speeds in reading ironic and literal target sentences.

As I show below, the psycholinguistic literature is dominated by at least three accounts that bicker over the priority of the literal meaning or the ironic meaning of a sentence. What is noticeably absent in this debate is that there is little focus on attitude ascription. It seems that the empirical debate does not take into account the main theories of Pragmatics nor the seminal experimental papers discussed in the previous section, which consider the communication of attitude the hallmark of irony. Let's take a look at the three theories that have dominated discussion in the Experimental literature.

3.3.1   The Standard Pragmatic Model

The Standard Pragmatic Model (SPM) evolved from the work of Grice (1975) and Searle (1979) and is one of the earliest and most influential approaches to the comprehension of figurative language. In the SPM one can see how the architecture of Grice's seminal work has been directly translated into a psychological explanation for understanding any linguistic phenomenon. At its simplest, the so-called SPM is a three-step process that involves 1) the computation of the semantic/literal meaning; 2) the recognition of a violation of a maxim; and 3) the computation of an implicature. However, it has been difficult to establish that these three steps actually occur (let alone in such an order) and, more fatally, such a three-step process seems too long and slow for explaining the rapid on-line pragmatic processing of an utterance. Studies like Schwoebel et al. (2000) and Dews & Winner (1999) are compatible with the Standard Pragmatic Model, but more recent accounts tend to argue against it. This has made it easy for critics to rail against the SPM and, in so doing, the entire Gricean approach. However, we would underline that Grice never intended his model to be used as a model of actual language processing. As pointed out in Noveck & Spotorno (in press) the way Gricean theory is transformed into "the SPM" is emblematic of a common pitfall in the Cognitive Sciences as underlined by David Marr (1982).

In his seminal work, Marr pointed out how one can advance theoretically at three different levels – often referred to as the *computational, algorithmic* and *implementational* levels of analysis – and how one can make progress by keeping the three separate and

complementary. We summarize the three quickly. The *computational* level makes explicit the input and output of the process as well as the constraints that would allow a specified problem to be solved. The *algorithmic* level describes how to get from input to output, and specifically determines which representations have to be used and which processes have to be employed in order to build and manipulate the representations. The *implementational* level provides a description of the physical system that should realize the process at, say, the neuronal level (see chapter 3 for more details).

It should be clear then that Grice's theory was designed at the computational level; the SPM was invented to practically mimic it at the algorithmic level. However, as Marr argued, it is neither necessary nor recommended to assume that the two need resemble each other. One can do theoretical work at the computational level without recourse to the algorithmic level and so on with any level with respect to the other two.

3.3.2 Direct Access View

More recently, Gibbs (1994, 2002) has offered an approach to figurative language comprehension that directly contrasts with the SPM. The only common element between Gibbs's approach, which has been called the Direct Access View, and SPM is that both make no distinction between irony and other forms of figurative language. In Gibbs's case, he suggests similar processing mechanisms for both figurative and literal language. According to this view, the comprehension of figurative language does not involve particular cognitive processes (Gibbs, 1994; Gibbs & Moise, 1997). This assumption is based on the notion that comprehending literal as well as non-literal meanings of a sentence largely depends on pragmatic knowledge and listeners' figurative modes of thought (Gibbs, 1994, 2002). Furthermore, Gibbs suggests that literal and non-literal meanings are not distinct from each other since they are both determined by contextual information. By use of pragmatic knowledge together with contextual information, the analysis of some aspects of word meaning should be sufficient for understanding intended figurative meanings (Gibbs, 1999, 2002). Thus, intended and contextually compatible meanings can be understood directly without leading to an incompatibility during semantic information processing. Therefore, the comprehension of sentences that achieves figurative meaning is proposed to be no more difficult than those that achieve literal meanings, since both meanings might be extracted from the foregoing context.

Evidence in favor of the Direct Access View comes from comprehension studies showing that latencies for comprehending literal and figurative readings of similar target-sentences are comparable. This is based primarily on a seminal study from Gibbs (1986) on sarcasm (see the next section for more details on Gibbs, 1986 and Gibbs, O'Brien & Doolittle, 1995 for further evidence in favor of the Direct Access View). More recent studies have reported that contextual information facilitates the recognition and comprehension of sentences conveying ironic meaning (e.g., Colston, 2002; Colston & O'Brien, 2000; Ivanko & Pexman, 2003) but that they still remain comparable to their literal counterparts.

3.3.3   The Graded Salience Hypothesis

The Standard Pragmatic Model and the Direct Access view can be considered two extremes of a spectrum that other proposals then fill. One of the most influential accounts that fills the gap is the Graded Salience Hypothesis, proposed by Rachel Giora (1997). Giora (1995) considers irony a form of indirect negation that relies on dissimilarity between the literal and implied meaning. According to the Graded Salience Hypothesis, the initial processing of lexical information is an encapsulated and graded process in which salient meanings of words or expressions are retrieved from the mental lexicon (Giora, 2003). During initial processing, contextual information is processed in parallel but neither interacts with lexical processes nor inhibits salient meanings when contextually incompatible (e.g., Giora, 2002; Peleg, Giora, & Fein, 2001). Salience is a function of properties such as familiarity, prototypicality and frequency and the meaning of a word, in order to be salient, has to be encoded into the mental lexicon. In case words or expressions have multiple meanings varying in salience, Giora (2003) suggests that this process is graded: more salient meanings are accessed earlier than less salient meanings. Thus, most salient meanings are always accessed initially irrespective of their literality or contextual support. This implies that the processing of figurative sentences only diverges from that of literal sentences during later phases of processing if accessed salient meanings cannot be integrated with contextual information. In that case the salient meanings have to give way to less salient but contextually appropriate meanings. As opposed to the Direct Access View, contextual information is proposed to have a very limited impact because they cannot restrict initial access of salient meanings that might be contextually incompatible.

Evidence for the Graded Salience Hypothesis comes from behavioral studies that investigate the comprehension of irony (e.g., Giora et al., 2007; Giora & Fein, 1999; Giora, Fein, & Schwartz, 1998). For example, Giora & Fein (1999) have shown that conventional ironies (e.g., "Very funny.") can be processed as easily as literal remarks, while unconventional ironies seem to require more effort. In addition, Giora et al. (e.g., Giora et al., 2007; Giora, 2011) presented evidence against the Direct Access View by showing several instances where reading an irony takes longer than reading a literal sentence, regardless of the amount of contextual information that should facilitate the interpretation of ironic remark.

3.4   The irony of the reading time data

From our Pragmatics-oriented point of view, it seems bizarre that the huge debate about irony processing has inspired multiple experimental studies but that it is practically devoid of proposals that consider attitude ascription in their descriptions of irony processing. We have three hypotheses about why this would be the case. First, all three -- SPM, the Direct Access View and the Graded Salience Hypothesis -- make no distinction between irony and other forms of figurative language. Second, all three focus on surface-level aspects of the linguistic stimulus without aiming to determine whether other resources and processes, namely ToM, play a role in the comprehension of an ironic remark. Third, the Direct Access View in particular, as well as the Graded Salience Hypothesis, are lexically based accounts. Arguably, these proposals might be strong in accounting for lexical-based phenomena such as metaphor – where the metaphorical vehicle needs to adjust to a target concept -- but they are arguably not as well disposed to describe irony.

To make this last point clearer, consider the well known metaphor "This lawyer is a shark." Here, the lexical concept SHARK needs to be adjusted in order to be applied to the target concept LAWYER (for a deeper explanation of the adjusted concepts see Sperber & Wilson, 2006; Wilson & Carston, 2008; Carston, 2002). However, irony is not this kind of phenomenon; the literal meaning of an ironic sentence does not need to be adjusted in order to derive the figurative interpretation; rather, the hearer has to use the literal meaning as a cue to go beyond the word(s) and grasp the ironic attitude.

While the thesis finds the lexically-based accounts interesting, it also aims to better identify how contributions from Pragmatics can impact the psycholinguistic debate about

irony. The thesis aims to make its contribution by trying to explain the mixed results that one finds in the experimental literature. In the next chapter we will explain how we tried to do so.

4. How to bring attitude ascription back to center stage

The studies that will be presented here have been designed to fill the gap between the theories of Pragmatics about irony, on the one hand, and the contemporary psycholinguistic literature, on the other. From our point of view, this means taking into account attitude ascription while one investigates irony processing. The bet was that paying attention to attitude ascription will allow us to account for the inconsistent results in the behavioral studies. Going in, we anticipated that attitude ascription can account for existing results in at least one of two ways. One is that participants could potentially acclimate to ironic readings through mindreading. The other is that different pools of participants could vary in their mindreading abilities in such a way that they can be more or less sensitive to the cues that anticipate an ironic remark. In the following, the thesis will show how both hypotheses were tested in three reading time experiments. In Experiment 1, the reading time of target sentences in stories were recorded when they represented either ironic or literal readings and while part of larger sets that included filler stories. Indeed, the experiment reports that target sentences in ironic versions are extraordinarily slow in the early parts of a session and become as fast as their literal counterparts by the end of the session. In Experiment 2, we aim to block the purported effect of habituation linked to irony by including decoy items among the filler items; these are stories containing negative events that are not followed up with ironic remarks. Finally, we investigate the Echoic Theory by manipulating the presence and absence of explicit to-be-echoed-primes of the ironies while the decoys remain. In all the Experiments, we ask participants to fill out a brief Autism-Spectrum Quotient (AQ) questionnaire (Baron-Cohen et al., 2001) to determine whether reading time performance can be correlated with mindreading ability. Experiment 1 reveals how individual differences can account for reading time slowdowns with respect to ironic utterances.

Shouldn't irony be effortful?

Nicola Spotorno[a], Ira A. Noveck[a,b]

[a] Laboratory L2C2, CNRS Université Lyon 1 FRE3406, 69675 Bron CEDEX, France

[b] Centre de recherche français à Jérusalem (CRFJ), MAEE-CNRS (UMIFRE 7) B.P. 547, 3 rue Shimshon, Baka, 91004 Jérusalem, Israel

SUBMITTED

Corresponding author:

Nicola Spotorno, L2C2 CNRS-Université Lyon 1, Institut des Sciences Cognitives, 67 Bd Pinel - 69675 Bron CEDEX, France.

Phone: +33 (0)4 37 91 12 71

E-mail address: nicola.spotorno@isc.cnrs.fr

Abstract

While some studies indicate that ironic -- as opposed to literal -- readings of utterances slow readers down, others indicate that the two are processed equally easily. We propose that mindreading processes are responsible for the mixed results as we present three experiments containing vignettes of which half use target-sentences having *Ironic* or *Literal* readings. Experiment 1 produced an *Early-Late* effect, which occurs when irony-readings specifically take longer to process in the first half of an experimental session. Experiment 2 replaced half the filler items with *decoys*, which are stories containing negative events (making them ideal for irony) that nevertheless lead to banal, non-ironic target-sentences; with their inclusion, one finds a main effect (showing ironic target-sentences taking longer to read than literally interpreted ones) and no *Early-Late* interaction. When primes to target sentences are explicitly included in all stories, including filler items and decoys, Experiment 3 reports that the *Early-Late* effect reappears, confirming expectations from Echoic Mention Theory (Sperber & Wilson, 1981). All participants filled out the Autism Quotient (AQ) questionnaire (Baron-Cohen et al., 2001) to determine whether reading times are correlated with mindreading. Experiment 1 revealed a significant positive correlation showing that Irony-minus-Literal reading-time differences in the second half of a session are positively correlated with the magnitude of autism-related social behavior.

Introduction

Is it more difficult to understand an utterance that is used ironically than when it is used literally and sincerely? To make this question concrete, consider an opera singer who utters (1) to her interlocutor after the two were part of a clearly awful performance:

(1)   "Tonight, we gave a superb performance."

Next, consider the same utterance after the two singers were part of a brilliant performance. Would the ironic interpretation (after the horrible performance) be understood as readily as the literal interpretation (after the astounding performance)?

On the one hand, it would make intuitive sense to suppose that an ironic meaning requires more effort than a literal one (Dews & Winner, 1999) . After all, irony takes the linguistically encoded message in order to access an interpretation about the speaker's intended meaning that goes beyond the meaning of its words while a literal interpretation does not appear to call for such further depth. To some extent, an *irony-is-effortful* view appears to be supported based on evidence showing that irony is linked with slowdowns. For example, in a reading time study, Schwoebel and colleagues (Schwoebel, Dews, Winner, & Srinivas, 2000) found that a sentence was read more slowly when it was meant ironically than when it was meant literally. Similarly, Filik and Moxey (2010) reported longer latencies for ironic readings over literal readings of the same target sentences in an eye-movement study. On the other hand, there is also evidence indicating that ironic utterances are integrated as easily as their literal counterparts (e.g., Gibbs, 1986; Ivanko & Pexman, 2003; Katz, Blasko, & Kazmerski, 2004). As we will describe in greater detail later, Gibbs (1986, Experiment 1), who spearheaded the Direct access view, also compared the reading times of sentences having the same surface form but in different contexts (e.g. "You are a fine friend" said so that it is either ironic or literally true) and he reported reading times that remained similar regardless of interpretation. While Ivanko and Pexman (2003, Experiment 3) showed how context can modulate the process of irony, they too found that -- at least under certain circumstances of a word-by-word reading task—latencies related to irony are equivalent to those that call for a literal interpretation.

Can the contradictory findings on irony ever be reconciled?

As Filik & Moxey (2010) recently pointed out, the evidence on (the primacy of literal readings in) irony-processing remains "mixed" (page 422). It appears then that one is forced to agree or disagree with findings, which is unproductive and weakens the psycholinguistic literature. Resolving that the science has not gone far enough, we set out to make sense of the discordant findings and by highlighting the largely overlooked role played by Theory of Mind.

In the remainder of the Introduction, we take the following three steps. First, we review the major pragmatic accounts of irony, starting with Grice, in order to highlight the prominent role played by attitude-ascription in these accounts. This review will show how early investigations, which were centrally concerned about attitude-ascription, have given way to the current debate about the extent to which literal-interpretations of ironic remarks are accessed during uptake. Second, we turn to the paradigms and findings from Gibbs (1986) because that paper is arguably most responsible for a) the irony-is-effortless view and for; b) providing support for one of the post-Gricean accounts, the echoic-mention theory (Sperber & Wilson, 1981; Wilson, 2009). While we point out that there are features in Gibbs's experiments that raise doubts about the conclusions drawn from them, we aim to separate these issues from those concerning the echoic-mention theory. Some of the features we highlight have been taken into consideration by others before us (e.g., Giora, 1995) while others have not. For example, we will highlight how negative events in Gibbs's study essentially telegraph that attitude-revealing remarks are looming, which could encourage a reader to anticipate an ironic remark. Finally, with the aim of uncovering Theory of Mind's role in irony processing, we describe how we modify Gibbs's paradigm. This leads us to present three experiments that essentially determine the extent to which one can isolate the role played by Theory of Mind in irony processing.

Theoretical background(s)

From the perspective of Paul Grice, figurative utterances are departures from a norm of literal truthfulness used to convey a related figurative meaning or *implicature*. Irony is said to be understood because the speaker is blatantly violating the first maxim of quality ("Do not say what you believe to be false"), which triggers implicatures that can be calculated from the literal meaning of the sentence. Although initially influential, the Gricean account is now increasingly in doubt due mostly to its emphasis on the presence of maxim-violations that are not always apparent. For example, consider the utterance: "I love sunny days!" when it is said

under a downpour. The utterance is ironic and is probably true as well, but there are no identifiable maxim violations.

The first challenge to Grice's approach came from the Echoic mention theory (Sperber & Wilson, 1981; Wilson, 2009; Jorgensen, Miller, & Sperber, 1984). According to this account, the speaker of irony is not expressing her own thoughts, but echoing a thought that can be attributed to some real or prototypical speaker, while expressing a dissociative (mocking, skeptical or contemptuous) attitude to that thought. When our fictive interlocutor says "I love sunny days!" on a cold and rainy day, what makes the utterance ironic is that the speaker is harking back to a possible and an appropriate remark for desired weather. As Wilson (2009, page 197) writes:

> … the point of irony is not to commit the speaker to the truth of the proposition expressed but, on the contrary, to express a certain type of derisory or dissociative attitude to a thought with a similar content that she attributes to some source other than herself at the current time. In other words, the speaker of irony is not expressing her own thoughts, but echoing a thought she attributes to someone else, and expressing her mocking, skeptical or contemptuous attitude to that thought."

It is through echoing, that the speaker makes her attitude apparent.

Clark and Gerrig (1984) proposed an alternative view which they called a "Pretense Theory of Irony." The main idea behind the Pretense account is that the speaker of an ironical utterance is not herself performing a speech act (e.g. making an assertion or asking a question) but pretending to perform one, in order to convey a mocking, skeptical or contemptuous attitude. This pretense mood would make an utterance ironic. With "I love sunny days," the speaker is pretending to be an unseeing person, perhaps a weather forecaster, exclaiming to an unknowing audience how much she likes beautiful weather. She intends the addressee to see through the pretense—in such rain she obviously could not be making the exclamation on her own behalf—and to see that she is thereby ridiculing the sort of person who would make such an exclamation (e.g., the weather forecaster). The addressee can take "delight" in "the secret intimacy" shared with the speaker in recognizing that ignorance (Clark & Gerrig, 1984, 122).

While the Echoic-mention and Pretense accounts each give primacy to the accessing of a speaker's attitude, they differ with respect to the object of the ironic remark. For Sperber & Wilson, it is some previously stated remark or shared cultural norm and for Clark & Gerrig,

it is practically a parody of the speaker who would have made such a remark. It should also be noted that while Grice focused on the blatant violation of truth in irony, he also took attitude into consideration for he recognized that a "hostile or derogatory judgment or a feeling such as indignation or contempt" had a role to play in understanding it (Grice, 1989, 53). Point is that all three of these seminal accounts agree that the communication of the speaker's attitude is crucial to the comprehension of irony.

A couple of studies tried to test predictions based on these seminal theoretical accounts (mainly the predictions of the Echoic Theory, e.g., Jorgensen et al., 1984; Keenan & Quigley, 1999; Happé, 1993; Gibbs, 1986). Most notably, Jorgensen, Miller & Sperber's (1984) investigation into Echoic Theory predicted that a sentence such as "Tonight we gave a superb performance!" would be easily interpreted as ironic if the story contains an antecedent (e.g. imagine the main speaker saying something like "Tonight we will be great!" earlier in the story) that is later echoed by the ironic statement. It also follows that participants would find the same target sentence harder to understand if the context does not include such an antecedent. The results of Jorgensen et al.'s (1984) questionnaire confirmed these predictions and provided the literature with the first set of data in favor of the Echoic Theory of irony. Later, Francesca Happé, in her seminal paper on pragmatic abilities in Autism Spectrum Disorders (1993), revealed that only autistic people with intact high level-mindreading abilities (namely the second-order Theory of Mind, which is the ability to interpret the thought of someone about the thought of someone else) are able to correctly interpret ironies. Happé's study provided strong evidence in favor of Sperber and Wilson's analysis of linguistic communication, which includes the Echoic account of irony under the umbrella of Relevance theory (e.g., Sperber & Wilson, 1986; Wilson & Sperber, 2012b).

Despite these early views on irony that clearly took into account concerns about Theory of Mind, a different debate arose in the psycholinguistic literature on the immediacy with which ironic interpretations are made when compared to literal readings. This is because Gibbs, in his processing paper, argued strongly against Grice's "Standard Pragmatic Model" by pointing out that there are a host of cases including indirect requests, certain metaphors and irony (and especially sarcasm) that give rise to the intended reading without requiring a contrast between the literal reading and a given situation that blatantly violates a maxim. Central to Gibbs's account, known as the Direct Access View, is the assumption that a figurative interpretation is constructed "directly" by the early integration of lexical and contextual information. When contextual information is sufficient, the figurative meaning is the first and only meaning that is activated; one does not need schedule a literal step during

the comprehension of a figurative utterance. In his own words, ''People need not first analyze the literal, pragmatic-free meaning of an utterance before determining its figurative, implicated interpretation'' (Gibbs, 1994, 421).

Before focusing on Gibbs's experiments, which serve as the basis for our own investigation, it is important to complete our diaporama on irony processing with the Graded Salience Hypothesis (Giora, 1997). This hypothesis emphasizes that the most salient meaning is the one that is accessed first regardless of whether it is literal or figurative. The salient meaning of a figurative remark is the one that is encoded in the mental lexicon of the audience. Salience is defined as a function of different features like familiarity, conventionality, frequency and prototypicality. The figurative interpretation of an utterance that is used very frequently to convey that meaning can be the most salient one of that utterance. Only in this case is it computed first during the comprehension process. In contrast if the figurative meaning is not encoded in the lexicon, more inferential steps -- and thus more cognitive effort -- will be necessary for arriving at the intended interpretation. Much recent work presents evidence as support for the Graded Salience Hypothesis (Filik & Moxey, 2010; Giora, 1997; Giora & Fein, 1999; Giora, Fein, Kaufman, Eisenberg, & Erez, 2009) and sets itself up as being in opposition to the Direct Access view.

As can be seen, the Direct Access versus Graded Salience debate differs markedly from the earlier one about attitude-ascription and its role in irony. The more recent debates are about the surface features of the words used in irony and their effect on the speed of processing. The net result is that the experimental literature has only scratched the surface concerning the role played by attitude ascription in irony. This turn is unfortunate because the idea that there is attitude ascription in some form remains uncontroversial. The main question is how does it figure into irony comprehension.

The time is ripe to reintroduce attitude ascription into the study of irony processing and for three reasons. First, the literature on mindreading (or Theory of Mind) has exploded in the cognitive sciences since the earliest experimental investigations have been published and there is an increasing wealth of resources on which to draw (for reviews, see Baron-Cohen, 1995; Leslie, 1987; Wellman, Cross, & Watson, 2001). Second, developmental work (Filippova and Astington, 2008) has actually shown a positive correlation between the development of the ability of understanding ironic remarks and advanced ToM abilities. The stronger children are at attributing mental states to a story character in classic ToM tests the stronger they are at correctly identifying intentions behind ironies (e.g., distinguishing between irony and deception). These classic ToM tests employ "second-order false belief

stories" (which examine children's awareness of a character's knowledge about another character's false belief; Astington, Pelletier, & Homer, 2002), ''strange stories'' (which tests children's awareness of a character's motivations to say or do something on the basis of his or her knowledge of another character's belief about the event described in the story; Happé, 1994), and ''faux pas stories'' (which examine children's awareness of a character's knowledge of another character's intentions, feelings, and motivations; Banerjee, 2000) Third, an investigation into attitude ascription could potentially disarm the standoff between the Graded Salience and Direct Access positions. Indeed, the aim of the present paper is to make sense of the literature's inconsistent findings by considering the hypothesis that attitude ascription in irony processing affects experimental outcomes in characteristic ways. As part of this effort, we now turn to the work that first considered the notion that irony could be effortless -- Gibbs (1986).

Gibbs's seminal investigation

Gibbs's (1986) experiments focused on the latencies of ironic and literal readings of target utterances. To produce different readings, contexts typically (and practically by necessity) employed a negative situation (e.g., a brother who absents himself from a task) to produce ironic interpretations (an interlocutor who says "You're a big help") and employed a positive situation (e.g. a collaborative friend who helps with a task) to justify using the same utterance literally (i.e., a context that renders "You're a big help" sincere). The seminal paradigm (i.e. Experiment 1) also included two control conditions. One used the negative context but followed it up with a brutally honest remark (e.g. "You're not helping me") while the second -- called the Positive/Compliment target condition -- used the positive context to follow it up with a banal remark (e.g., "Thanks for your help"). The results showed (a) that the latencies for ironic readings were shorter than those for brutally honest ones; (b) that the reading times for ironies were comparable to "literal" ones and; (c) that the banal positive comment prompted the fastest reading times of all. It was result (b) that justified Gibbs in concluding that ironic statements are as easy to process as literal ones. However, note that both were slower than the banal remark and faster than the brutally honest one.

In a second experiment, Gibbs went further and linked his claims to Sperber & Wilson's Echoic-mention account. He had ironic statements serve as concluding remarks in two sorts of (slightly different) contexts. In one, there was an explicit mention of a feature that would serve as the basis of an ironic remark. For example, in the *Echoic* condition, the reader is told about Gus who saw an advertisement about joining the Navy (while making reference

to a well known advertising campaign of the time whose tag line was *not just a job, but an adventure*); after joining the Navy and seeing how uneventful his job really was, Gus says "This sure is an exciting life." In the Non-echoic version of the story, the sentences containing the explicit mention of the advertisement are excised. Gibbs found reading time facilitation on the ironic-target item when there had been an explicit mention (of the advertisement), i.e. when there was a basis in the story for an echoic mention later, with respect to cases that had no prior explicit mention of the thought that was echoed in the character's remark. The data from this experiment served as the basis for claiming that echoic-mention facilitates irony processing.

It might not be obvious that our summary has something to do with mindreading until we highlight three features of the paradigm that help readers anticipate outcomes. First, note how negative contexts (e.g., a brother who does not show up to a *rendez-vous*) are consistently linked with a strong attitude-revealing utterance -- either in the form of an ironic remark (e.g., "you're a big help") or a brutally honest one (e.g. "You're not helping me") -- while positive contexts are associated with less provocative utterances that are expressed sincerely (in the case of the literal reading of "you're a big help") or with banal continuations ("Thanks for your help"). When some sort of nastiness prompts the expectation that the bereaved will soon share a provocative thought, readers will arguably be anticipating it over the course of an experimental session. Such telegraphing could conceivably lighten the effort needed to carry out ironic interpretations bringing their reading times down when compared to unanticipated cases.

Second, one would think that a Literal condition would lead to floor effects, but this was evidently not the case since it was the banal continuations (the target sentences in the *Positive/Compliment* condition) that were read significantly faster than all others. It is thus not clear that target sentences in the "Literal" condition were read as a control nor as innocently as intended. This could be due to the fact that there were just 16 stories of which half were negative and whose endings were provocative and attitude-revealing; this could lend the Literal condition's target sentences to provocative interpretations as well. In other words, when every other story recounts a miscommunication or a disaster followed by a strong remark, even a sincere "you're a fine friend" might transmit a double entendre. If we are right, this indicates the extent to which mindreading is prevalent in reading (even literally interpreted) texts. After all, literal texts were significantly slower than banal follow-up statements.

A third way to see an aspect of mindreading is with respect to (a) the way story-construction varies across Experiments and (b) the investigation of echoic mention. In terms of (a), one collection of vignettes (in Experiment 1) presents target-items that are remarks made by one character to another in light of an interaction between the two whereas another collection (in Experiment 2 on echoic-mention) presents target-items that make reference to a situation external to both interlocutors. The manipulation in Experiment 2 was carried out to address considerations from Clark and Carlson (1982), but the upshot is that the focus of ironic utterances was not consistent across the paper's experiments. This non-standardization of materials prevents one from appreciating how new manipulations (e.g. echoic mention in Experiment 2) compare with outcomes from previous experiments. In a similar vein (concerning (b)), stories in the "non-echoic mention" condition in Experiment 2 are two sentences shorter than others in the same experiment. This raises the possibility that the apparent facilitation linked to the "echoic mention" condition (albeit with different sorts of stories) could be due to a participant's increased readiness for the ironic remark. If an ironic remark ends one story after four lines while many stories in the same session are as long as seven, it could be that the lack of facilitation is due to a lack of preparation for irony in the "non-echoic" stories (and heightened expectations for irony in the explicitly Echoic ones).

Uncovering a role for attitude ascription in experimental tasks

From the point of view of mindreading, these issues point to the possibility that attitude ascription plays an essential role in irony-processing. Also, they indicate that attitude ascription can be isolated experimentally. How do we intend to do that? One way is to test the hypothesis that participants habituate to ironic attitudes over the course of an experimental session. We thus propose that attitude ascription is effort demanding, especially as a one-off event but that an anticipated attitude, e.g. one that reappears under specific circumstances, facilitates mindreading and thus makes irony appear non-demanding. If this occurs over the course of an experimental session one should be able to note that.

With this view, one ought to find an interaction in which ironic readings start out distinctly slower from literal readings early on in a session and become progressively faster later, perhaps as fast as utterances that call for literal readings. Our prediction is based on the hypothesis that one can acclimate to ironic readings through mindreading whereas literal readings ought to represent a lower bar that does not as readily rely on it. It is such habituation that would account for the "mixed" results in the literature because one can imagine that stimuli in irony experiments vary. A set of stimuli that well anticipates ironical readings for

participants could make ironies read as fast as literal ones even early on in an experimental session. Materials that make irony anticiaption more unpredictable ought to delay habituation.

A second way to account for the literature's incongruent results is to assume that the application of a mindreading ability varies across subject pools. Perhaps a majority of participants in one subject pool is especially adept at mindreading in one University and is not in another pool. In order to address the potential variability of Theory of Mind abilities, we will differentiate among participants and their capacity to mindread. This hypothesis is motivated by links found previously between pragmatic enrichments of scalar utterances and social ability scales (Nieuwland, Ditman, & Kuperberg, 2010) using Baron-Cohen and colleagues' Autism-Spectrum Quotient (AQ) (Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001). If Nieuwland et al.'s findings are generalizable, it would make sense to investigate readers' AQ with respect to irony, which is arguably among the most recognizable pragmatic forms in the human repertoire.

In the next section of the paper, we present three Experiments. In Experiment 1, where we describe how we standardize story-construction, we compare the reading time of target sentences (having either ironic or literal readings) while including filler items. To anticipate, this study produces what we dub an *Early-Late* effect which occurs when irony-readings, specifically, take an extraordinarily long time to process in the early part of (the first half of) the experimental session when compared to literal ones. In Experiment 2, we aim to forestall the Early-Late effect linked to irony by including *decoy* items among the filler items; these are stories containing negative events that are not followed up with ironic remarks. This manipulation demonstrates that the Early-Late interaction relies on linking negative events with ironic follow-ups over the course of the Experimental session. Finally, we investigate the Echoic mention theory by manipulating the presence and absence of explicit primes of the to-be-echoed ironies. In this experiment, which includes decoys, the Early-Late effect reappears. In all Experiments, we ask participants to fill out a brief AQ questionnaire to determine whether reading time performance can be correlated with a mindreading ability.

Experiment 1

The aim of this Experiment is to test whether one can find evidence showing that participants habituate specifically to an ironic attitude over the course of an experimental session. In order to accomplish this, we prepared vignettes whose context can be minimally modified so that each target sentence can provide either ironic or literal readings. The critical vignettes were mixed with fillers which were plain stories without ironic remarks. In order to investigate

irony processing in depth, we a) take into account the moment at which any given stimulus was presented (i.e., we compare the reading times for the stimuli presented in the first half of the experiment to those of the second half), and we; b) evaluate whether or not individual differences (with respect to a participant's AQ) are linked with irony comprehension. With respect to (a), our prediction is that the reading times for irony drop off at a rate that is steeper than that for literal readings over the course of the experimental session. With respect to (b), our prediction is that participants with scores that tend upward (high scores on the scale indicate greater autistic behavior) will generally need more time to understand ironic utterances throughout the session.

In order to standardize the stimuli, we modify Gibbs's paradigm in five ways in order to address the concerns described in the Introduction. First, we use stories that always consider an event that occurs to both of two interlocutors (rather than have the aggrieved make a remark to a second character who is the source of the grief). Second, all vignettes are seven lines long. Third, only original expressions of irony are used (conventional expressions such as "fine friend" are avoided). Fourth, irony is detectable through the last word (which is possible in French) so as to avoid telegraphing. Fifth, we systematically give Baron-Cohen's AQ to all participants. This approach will be adopted for all experiments in the paper.

2. Materials and methods

2.1 Participants

Forty-three native speakers of French participated in Experiment 1. They were recruited from the *Université de Lyon* (Lyon 1 and Lyon 2). Their ages ranged from 18 to 27, with a mean of 21. Each participant reported to have normal or corrected-to-normal vision.

2.2 Materials

20 sentences that can be interpreted as ironic or as literal were created. For each sentence two contexts were written: one encouraging an ironic interpretation and the other a literal one. In this way, a set of 20 contexts (that can be, with slight changes, the source of either an ironic or literal target sentence) was set up (see Table 1). Every story has the same structure:

I. Two characters interact during a daily situation. Very close relationships between characters (such as close friends or couples) were avoided.

II. Each story is made up of seven lines. Each of them has a maximum length of 91 characters (spaces included).

III. The first three sentences introduce the two characters and the situation.

IV. The fourth and fifth sentences describe the development of the situation, which can be positive (in the literal version) or negative (in the ironic version).

V. The sixth line is the target sentence. The lengths of all the target sentences are between 10 and 12 syllables.

VI. The seventh line is a wrap-up sentence that provides a sensible conclusion to the story. Its length is always between 18 and 20 syllables (see table 1 for an example).

VII. Differences that distinguish ironic and literal versions of each story are due to changes made to the fourth and fifth lines only.

Insert Table 1 about here

Twenty fillers were also written. They respect the same structure as the literal versions of the stories and they follow the same constraints, except that the sixth line is neither a literal reading nor an ironic one. Fillers are best viewed as continuations (see Table 1 for an example and Appendix C for further examples). A yes/no comprehension question that pertains to some general detail other than the target sentence of the story was prepared for all stories. The primary purpose of the questions was to make sure that the participants paid attention while reading the entire story. The correct answer was "yes" for half of the questions and "no" for the other half.[3]

A French translation of the Autism-Spectrum Quotient questionnaire (Baron-Cohen et al., 2001), known as the AQ, was used for testing participants' social/pragmatics abilities. This questionnaire was designed to provide an indirect measure of a reader's pragmatic abilities on a continuum from Typical to Autistic. The AQ is divided into 5 subscales: *social skill*, *attention switching*, *attention to detail*, *communication*, and *imagination*. Given Nieuwland et al.'s (2010) findings, we are especially interested in the *social skill* subscale. Some examples of the items from this subscale are: "I prefer to do things with others rather than on my own", "I find social situations easy" and "I would rather go to a library than a

---

[3] Experiment 2 will describe a rating study that confirms that the ironic and literal target sentences here were understood as intended. The rating study is described in Experiment 2 because it also evaluated stories from a third condition used there.

party". Each participant filled out the Autism-Spectrum Quotient (AQ) questionnaire and the score was used to uncover associations with observed reading times.

2.3 Procedure

The participant was seated in front of a laptop PC running *Presentation* software (Neurobehavioral Systems, www.neurobs.com). Before starting the test, a training session of two stories was performed in order to familiarize participants with the experimental procedure. Each story was presented line by line in a self-paced manner. At the end of the story the comprehension question was presented and the participant was required to answer "yes" or "no" by pressing one of two buttons. Each line was left-justified and presented at the center on the computer screen in white, 24-point font against a black background. Each participant saw 10 ironic stories, 10 literal stories and 20 filler stories. The order of the presentation was pseudo-randomized so that 5 ironic stories, 5 literal stories and 10 fillers were presented in each half of the experiment. This constraint was introduced to allow for a balanced analysis on the basis of the order of presentation of the stimuli (see the paragraph "Analysis" below for details). Participants were instructed to read for comprehension and to answer the questions naturally. After the experimental session, each participant filled out the AQ questionnaire using the same computer as the self-paced experiment.

Insert Figure 1 about here

----

3. Results

3.1 Analysis

The answers to the questions were analyzed separately in order to verify that participants maintained a satisfactory level of attention. Only those who answered at least 70% of the questions correctly were included in the analysis. With this criterion, two participants were excluded from the analysis. We focus our statistical analysis on the reading times of target sentence (line 6), but we also report analyses of reading times concerning the wrap-up sentence (line 7). For each, reading times that were 2.5 standard deviations above the mean were considered outliers as were reading times inferior to 900 msec. (which is what we considered to be the absolute minimum needed to read through a sentence). These constraints led us to reject 8.8% of the data.

The reading times were log transformed to improve the conformity of the data to the standard assumptions of ANOVA (e.g., Howell, 1997). Repeated measures ANOVAs are performed using the variable *Type of story* (*Ironic* or *Literal*) as an independent variable. The data were also split according to the session-half -- "first half" or "second half" -- in which a story appeared. We refer to this distinction as *Early* and *Late*. Analyses were focused on the Irony versus Literal conditions since these were designed for comparisons from the start. However, we do statistically characterize filler items in order to complete the picture. All analyses included both participants and stimuli items as random effects in our model. Item analyses involved summing over all participants but distinguished between the two types of story. By convention, we refer to F-values obtained with participants as the random factor as F_1, while F-values obtained with items as the random factor as F_2. The score in the AQ questionnaire was analyzed separately and then simple regressions were applied in order to identify relevant correlations between the reading times of the target conditions and the AQ scores.

3.2 Target sentence

Figure 2 shows the reading times of the Irony and Literal conditions broken down into the Early and Late parts of the Session. One can see that reading times speed up over the course of the experiment overall. However, the target sentences in the *Ironic* condition were read more slowly (1981 +/- 70 msec.) overall than in the *Literal* condition (1704 +/- 52 msec.) and this is partly due to the fact that ironic readings took longest to process in the early part of the session. We also point out that the reading time of filler items pattern with the literal sentences (whose reading time mean is 1740 +/- 66 msec.).

An ANOVA with Log(reaction time) as the dependent variable, and *Type of Story* (Ironic / Literal) and *Part of Session* (Early/Late) as within-subject factors was carried out showing that the difference between the *Ironic* and *Literal* conditions is reliable, $F1 (1,39)= 29.85, p < .001, F2 (1,19) = 9.63, p < .01$, as is the main effect for Early-Late, $F1 (1,39)= 15.07, p < .001, F2 (1,19) = 17.98, p < .001$. Critically, the interaction between *Type of story* and *Part of Session* is also reliable, $F1 (1,39) = 5.59, p < .05, F2 (1,19) = 10.74, p < .005$. This confirms that ironic-readings take longer than literal readings when they first appear as part of the experiment and that they become comparable to literal readings in the second half.

3.3 Wrap-up sentence

The wrap-up sentence provides further information about irony processing since the follow-up sentences were also identical across the main conditions and of comparable length throughout. Here, we find that the reading times remain significantly longer in the *Ironic* condition (2692 +/- 104 msec.) than in the *Literal* condition (2423 +/- 94 msec.) $F1$ (1,39) = 36.69, $p < .001$, $F2$ (1,19) = 8.7, $p < .01$. Likewise, there is a main effect for Early and Late halves of the session, $F1$ (1,39) = 13.27, $p < .001$, $F2$ (1,19) = 10.5, $p < .005$. Unlike for the target sentences, the interaction between *Type of story* x *Part of Session* is not significant $F1$ (1,39) = .02, $p > .1$, $F2$ (1,19) = .1, $p > .1$. This indicates that the reading times of the wrap up sentences associated with the *Ironic* condition are generally effort demanding when compared to the *Literal* condition.

Insert Figure 2 about here

3.4 Correlations with AQ

While using the same distinctions for reading times (*Type of Story* and *Part of Session*), we investigated a limited set of correlations between AQ scores and the experimental conditions. These included investigations of participants' general AQ score (and subscale AQ score) and their mean target-sentence reading time in the *Irony* condition as well as the mean difference between ironic and literal readings for each participant (remember that each participant was given a unique mix of ironic and literal versions of the 20 stories). We report a significant positive correlation between the *social skill* subscale (the one employed in Neiuwland et al., 2010) and the *difference in reading times* for the target sentence between the *Ironic* and *Literal* conditions in the second half of the experiment, R = 0.326, $p < 0.05$. This indicates that the level at which a reader maintains a difference in reading time between the Irony and Literal conditions in the second half of the Experiment is positively correlated with those who score high on the social skill subscale. To put it another way, those who speed up on ironies over the course of the Experiment so that their readings become comparable to the literal cases are associated with scoring low on the scale (which means greater relative sociability compared to those high on the scale).

Insert Figure 3 about here
----
4. Discussion

The experiment was designed to determine whether the comprehension of irony requires extra effort. Several findings confirm that it does. This is seen most clearly in the target sentences which, when read ironically, take longer to advance than sentences whose meanings are closer to the literal. However, as we anticipated, the details tell a slightly more complex story. Reading time slowdowns linked to ironic target sentences are especially prevalent in the first half of the experiment; when looking at the second half, differences tend to disappear. This supports our hypothesis that participants habituate to the ironic attitude in this task over the course of an experimental session.

As with pragmatic interpretation generally, participants have to go beyond the linguistic code to understand a speaker's intended meaning. In the case of irony, this entails interpreting the speaker's attitude. While there is a minimal amount of time needed to read through a sentence's linguistic meaning, capturing an attitude might have its own trajectory. The interaction we report is arguably due to the reader's focus on the attitude ascription associated with irony, which is not too difficult once one is made aware of it and made aware of when it is due (in the target sentence). Interestingly, reading times for the wrap-up sentences, which were identical in both the *Ironic* and *Literal* conditions, were longer in the Ironic condition throughout the entire experiment. This indicates that the difference in terms of reading time in the target sentence is long-lasting even if one can habituate to an ironical attitude. We had no precise predictions about this post-ironic effect, but it may be due to the interference from extra cognitive effects that an ironic remark engenders (e.g., emotional effects) but that a literal sentence does not.

The habituation effect to the ironic remark is an intriguing finding. It is also associated with a positive correlation between the subscale *social skill* of AQ and the difference between the reading times in the *Ironic* and *Literal* conditions in the second half of the experiment. Those who score high (and who are closer to the autism spectrum) are more likely to *not* habituate. It seems then that participants' predisposition in social interactions is correlated with the way one potentially habituates to an ironic attitude.

Given the role of attitudes in appreciating ironic stories, its suspected role in habituation, and the way individual differences are implicated, we hypothesize that this tendency to habituate can be – to some extent – discouraged (at least among those who readily adapt to the task). How can an experiment prevent a participant from habituating to irony? As we pointed out earlier, irony is typically investigated by presenting a positive remark in the wake of a negative situation. This is how stimuli are designed in the great majority of the studies on irony, including the present one. In the literal stories, on the other hand, the plots

are positive throughout. Moreover, fillers are usually close to the literal condition in that they are plain, positive stories without an ironic remark. The upshot is that ironic stories are usually the only ones with a negative context which could, in effect, become a cue for an upcoming ironic remark. This is why the next study aims to lower the probability that irony is to be predicted in the wake of a negative context. We accomplish this with what we call *decoy* stories that start out with a negative event (e.g., an awful performance) and then follow up with a banal continuation (e.g., "We'll do better next time"). This should prevent a reader from viewing negative contexts as cues for upcoming ironic remarks.

Experiment 2

One of the paradigm's features of Experiment 1 is that ironic remarks consistently follow a negative situation. In order to determine whether this plays a role in the habituation effects we reported there, we replace ten of twenty fillers with stories containing negative events not having ironic endings. These decoys have the added advantage of balancing the polarity of the stories in the design (from one quarter being negative to half). Our aim is to prevent negative events from being cues to readers that an ironic attitude is looming. This should help keep ironic interpretations spontaneous throughout the task and to eliminate the *Type of story * Early-Late* interaction reported above. The question is whether it will block the Early-Late habituation effects linked to ironic readings.

2. Method

2.1 Participants

Twenty-four native speakers of French participated in Experiment 2. They were recruited from the *Université de Lyon* (Lyon 1 and Lyon 2). Their ages ranged from 18 to 26, with a mean of 21. Each participant reported to have normal or corrected-to-normal vision.

2.2 Materials

The pool of critical items was the same as the one in Experiment 1. There were 20 contexts that laid the foundation for either ironic or literal target sentences. Also, 10 fillers were identical to 10 fillers of the first experiment. However, another 10 were replaced with 10 decoys. Decoys were 7-sentence-long stories in which a negative event led to a banal utterance (for an example, see Table 2 and appendix B). In this new control condition, as in the other conditions, the length of sentence 6 is between 10 and 12 syllables, and the length of sentence 7 is between 18 and 20 syllables.

A rating study was conducted to verify that the specific stimuli used in this experiment (as well as the last one) were perceived as intended. Twenty-six participants (13 women), whose ages ranged from 19 to 35 (with a mean of 27) and who did not participate in any reading time experiment, were asked to read 50 stories (2 from each of the frameworks plus 10 that concern *decoys*) and to rate the extent to which a story's target sentence was ironic on a scale from 1 (not at all ironic) to 5 (very ironic). Whereas the 40 *Ironic* and *Literal* stories were pseudorandomized and balanced across two lists, all 10 decoys were included for each participant. Ironic target sentences were rated as highly ironic (mean of 4.5), while literal sentences and the banal lines from the decoy stories were rated as low on the ironic scale (1.2 and 1.4, respectively). Repeated measure ANOVAs showed significant differences between (i) the *Ironic* and *Literal* conditions and (ii) the *Ironic* condition and *Decoys* (both at $p < .001$, corrected for multiple comparisons using the Tukey method). The comparison between the *Literal* condition and the *Decoys* was not significant ($p = .1$)

Insert Table 2 about here

2.3 Procedure

The procedure was identical to Experiment 1's. At the end of the experimental session each participant filled out the AQ questionnaire.

3. Results

3.1 Analysis

We follow the same procedure as Experiment 1. This time, 6.7 % of the data were considered outliers. In this case, no participant was excluded from the analysis due to her percentage of wrong answers. As in Experiment 1, we first summarize the data from the Target sentences, before turning to the wrap-up sentences and (this time, the lack of) AQ score correlations.

3.2 Target sentence

As in Experiment 1, the analysis focused on the *Ironic* and *Literal* conditions. As one can see in Figure 2, the *Ironic* condition's target sentence was read more slowly (2141 +/- 118 msec.) than the *Literal* condition's (1828 +/- 90 msec.) and there appears to be habituation. However, there does not appear to be an interaction. A repeated measure ANOVA confirms these observations. The difference between the Ironic and Literal target sentences is statistically

significant, $F1$ (1,23)= 25.83, $p < .001$, $F2$ (1,19) = 7.66, $p < .01$ as is the Early-Late effect $F1$ (1,23)= 50.08, $p < .001$, $F2$ (1,19) = 26.20, $p < .001$. There is no *Type of story * Part of session* interaction, $F1$ (1,39) = 1.56, $p > .2$, $F2$ (1,19) = .06, $p > .5$ (see Figure 4). As in Experiment 1, we note that the reading times of the fillers pattern with the literal sentences more than they do to with the ironic target sentences (decoys: 1746 +/-94 msec.; positive fillers: 1819 +/- 97 msec.).

Insert Figure 4 about here

3.3 <u>Wrap-up sentence</u>

Overall the reading times were significantly longer in the *Ironic* condition (3019 msec. +/- 203 msec.) than in the *Literal* condition. The repeated measures 2 (Type of story: *Irony vs. Literal)* x 2 (Part of session: *Early-late)* ANOVA showed main effects for both Type of story $F1$ (1,23) = 18.91, $p < .001$, $F2$ (1,19) = 9.08, $p < .01$ and Part of session $F1$ (1,23) = 8.07, $p < .01$, $F2$ (1,19) = 5.26, $p < .05$. The interaction between the two was significant only when using participants as a random variable $F1$ (1,23) = 4.23, $p = .05$, $F2$ (1,19) =0.34, $p > .5$.

3.4 <u>Correlations with AQ</u>

No relevant correlation was found between AQ scores (general or subscales) and reading times.

4. Discussion

The replacement of ten positive fillers with ten decoys led only to main effects, i.e. a non-interaction, with respect to the *Type of story* and the *Part of Session* variable. The design of Experiment 2 masks the anticipation of ironies better than Experiment 1 and the lack of an interaction occurred among the target lines arguably because a very reliable cue (the one-to-one relationship between negative contexts and ironies) had been mitigated. As a result, participants were less able to anticipate attitude-ascription in the study. While the results of the present experiment are in line with our prediction they also underline how experimental bias was potentially present in prior experiments. These data show how features such as the consistent coupling of negative occurrences with ironic remarks, which we employed in Experiment 1, facilitate participants' on-line reading of irony, especially towards the end of an experimental session.

Unlike the results from Experiment 1, which indicated that readers vary with respect to their ability to anticipate attitudes and in a way that varies with their social score on the AQ, there are no significant correlations to report with respect to habituation. The decoys in Experiment 2 have apparently prevented the more pragmatically talented from anticipating an ironic attitude. We consider this further evidence that decoys had their desired effect.

Now that we established that decoys prevent participants from anticipating ironies, we now turn to Experiment 3, which is designed to specifically investigate the Echoic Mention theory. We are motivated to investigate this theory for two reasons. One is that while we have employed Gibbs's design while not finding direct support for the direct access view, we want to determine whether other related claims from Gibbs are equally affected. The other concerns our immediate experimental claims about the role of mindreading habituation over the course of a session. Now that we have established that decoys can block mindreading habituation, would the explicit presence of a what-we-will-refer-to as a *prime* facilitate irony comprehension and either eliminate differences between ironic and literal reading times or bring back the habituation reported in Experiment 1? In other words, following Sperber and Wilson's proposal, we investigate whether an explicit mention of a thought in a text can facilitate the comprehension of, or at least prompt habituation to, ironic remarks (i.e. these are unlike the cases of irony in Experiments 1 and 2 where the echo refers to something that is only implicit in the stories).

Experiment 3 employed the same stories as the previous experiment, but we modified a line in each story so that it could serve as a prime for the target line. Likewise, decoys remained but they too were slightly altered so that they more closely resembled the ironic stories. We predict that an explicit prime ought to be easier to echo than an implicit one. We anticipate that the presence of primes facilitates the echoic aspect of irony understanding, thus simplifying attitude ascription in irony. In the event that an explicit prime is as effective as Gibbs claimed, it should prompt an equivalence in reading times between Literal and Ironic target sentences. Another, less radical possibility is that the habituation effect will reappear. Such an outcome would indicate that participants take advantage of cues that can facilitate irony comprehension but that there is still some effort-demanding attitudinal work to be done early on in the session. Of course, and especially given the presence of decoys, the addition of an explicit prime could have no effect at all on irony comprehension.

This is a useful test of the echoic-mention theory because it assumes that a prime is an integral part of irony comprehension. Compared to the results of Experiment 2, to which this study is identical in nearly every way but with respect to the explicit mention of the prime,

one should find some form of facilitation for irony-comprehension. It is unclear whether one will find correlations between pragmatic abilities and irony comprehension in this case because the cue's explicit nature obviates the need for the sort of pragmatic skill that is found among socially inclined individuals. This is unlike Experiment 1 where the cue to the anticipation of irony was more covert and required that participants be attentive throughout the experimental session.

Experiment 3

Imagine that the opera singer who uttered (1) in our opening story explicitly indicates that she has high expectations; in such a case, the utterance "Tonight, we gave a superb performance" after an awful show would arguably sound like an ironic remark even more immediately than in our earlier versions because it echoes the singer's hopeful prediction made earlier. From the point of view of Echoic Mention Theory, this example exposes even more clearly how irony works because an ironic remark generally is a sentence that retrieves a thought while expressing a mocking, skeptical or contemptuous attitude toward it. If the echoed thought had been expressed overtly, the comprehension of the link with the ironic sentence should be recognized more easily. We argue that the link between the prime and the ironic sentence should act as the extra cue that facilitates ascribing a mocking, skeptical or contemptuous attitude to the speaker. The question for us concerns the strength of that effect. Will it make ironic comprehension appear as easy as its literal counterpart? If so, will this occur from the earliest moments of an experimental session?

2. Method

2.1 Participants

Twenty-four native speakers of French participated in Experiment 3. They were recruited from the University of Lyon (Lyon 1 and Lyon 2). Their ages ranged from 18 to 25, with a mean of 20. Each participant reported to have normal or corrected-to-normal vision.

2.2 Materials and Procedure

The pool of critical stories was the same as that of Experiment 2 (fillers and decoys included), but in each story the third sentence was replaced by an explicit prime. The presence of the prime is the only factor that distinguishes it from Experiment 2 (see Table 3 and the appendixes). The procedure was exactly the same as it was for the prior experiments.

3. Results

3.1 Analysis

Analyses were carried out in the same way as those from the previous experiments. In this case, roughly 7 % of the data were excised because the reading times were over 2.5 standard deviations longer than the mean or under 900 msec. As in Experiment 2, no participant was removed for having provided a disproportionate percentage of wrong answers to the comprehension questions.

3.2 Target sentence

Figure 5 shows the reading times of the target sentences in the Irony and Literal conditions broken down into the Early and Late parts of experimental sessions. One can see that the target sentences in the *Ironic* condition were read more slowly (1991 +/- 86 msec.) than in the *Literal* condition (1668 +/- 54 msec.) and that the Ironic readings took longest to process in the early part of the session (early part = 2286 +/- 132 msec.; late part = 1684 +/- 91 msec.). The repeated measures 2 *Types of story* x 2 *Parts of Session* ANOVA showed that the main effects for Type of story $F1$ (1,19) = 15.47, $p < .001$, $F2$ (1,19) = 17.99, $p < .001$ and Part of Session $F1$ (1,19) = 13.29, $p < .005$, $F2$ (1,19) = 30.92, $p < .001$ were reliable as was the *Type of story * Early-late* interaction, $F1$ (1,19) = 11.43, $p < .05$, $F2$ (1,19) = 14.26, $p < .05$ (see Figure 5).

As we did for the prior experiments, we point out how the reading times of decoy items pattern with the experimental sentences. The sentences that appear in the target slots in the decoy items pattern with the Literal sentences (1748 +/- 75 msec.). However, in this case, the speed of the positive filler items (1854 +/- 82 msec.) is intermediate between the literal and ironic target sentences.

3.3 Wrap-up sentence

The reading times in the *Ironic* condition (2800 +/- 117 msec.) were longer than in the *Literal* one (2497 +/- 112 msec.). The repeated measures ANOVA 2 *Type of story* x 2 *Early-late* showed main effects of the *Type of story* by participants only $F1$ (1,20) = 6.64, $p < .05$, $F2$ (1,19) = 3.69, $p = .07$, and main effects for *Part of Session* but only using subjects as random variable $F1$ (1,20) = 8.28, $p < .01$, $F2$ (1,19) = 2.71, $p > .1$. The interaction between the two conditions was not significant $F1$ (1,20) = .6, $p > .2$, $F2$ (1,19) = .01, $p > .5$.

Insert Figure 5 about here

3.4 Correlations with AQ

No relevant correlation was found between AQ score (general or subscales) and the reading times.

4. Discussion

The explicit mention of a prime that can be echoed by the target utterance had a clear effect on the comprehension of irony as can be seen by the presence of the interaction between the variables *Type of story* and *Part of Session*, an interaction similar to the one reported in Experiment 1. This occurs despite the presence of decoys, which prevented the interaction in Experiment 2 when there were no explicit primes. This is in line with predictions from the Echoic Mention account and is consistent with the findings from Gibbs (1986), even though the materials there produced null effects between the ironic and literal conditions. The Echoic Mention Theory argues that the irony comprehension process is made possible by the reference that the ironic utterance makes to a target thought. In cognitive terms, the reference can be translated in the link between the processing of the ironic sentence and the thought elicited by the *to-be-echoed* prime and this link facilitates the ascription of an attitude to the speaker. We thus link this effect with attitude ascription and specifically with Sperber and Wilson's Echoic-Mention Theory.

Based on talks we have given, we have heard potential criticisms suggesting that the facilitation is due, not to Echoic Mention, but to a more simple lexical priming. That is, it is argued that participants reading times are facilitated because they have already heard a line similar to the one stated in the prime. However, there are two factors that argue against this. First, the prime and the target sentence were never the same sentence in each story, so it is hard to argue for something along the lines of repetition effects. Participants are not rereading the same lines. A look at the materials reveals that the new lines added refer to the same event but never using the same set of words. Second, the prime was introduced in all the stories of every condition (fillers and decoys included). Thus, if it were the case that lexical priming is responsible for speedups, it would arguably occur throughout the task and to all conditions. Such an outcome would not be consistent with the interactions we report here. Given the interaction, we conclude that the *to-be-echoed* prime has a specific effect on irony processing.

General discussion

We opened the paper by asking whether it is more difficult to understand an utterance expressing irony compared to the same remark used sincerely and literally. We described how the data from the literature is currently mixed on this question and that existing psycholinguistic theories have arguably run out of ways to address it. Our strategy was slightly different. In looking for a way to account for the mixed results, we considered attitude-ascription, a central feature of pragmatic theories on irony that has largely been ignored in psycholinguistic analyses, as potential cause for the inconsistencies. This calls for characterizing mindreading's role in reading times.

We carried out three studies that ultimately show how attitude ascription can, under specific (though unexceptional) experimental conditions, eventually trump linguistic decoding to the point that ironic readings can appear as fast as literal readings of the same target sentences by the end of a session. When there was a one-to-one relation between a negative event and an ironic remark (Experiment 1), ironic remarks took longer to process than literal ones early on in an experimental session only (the first half). In other words, one finds null effects in the second half of sessions. When that one-to-one relation between negative events and ironic remarks was halved through the introduction of decoys (Experiment 2) one finds that the above interaction disappears, leaving a main effect showing that ironic readings are generally more effortful than literal ones. This shows that one can maintain slowdowns for ironic utterances (when compared to their literal counterparts) as long as they appear to remain spontaneous. Inspired by Echoic Mention Theory (Sperber & Wilson 1981; Wilson, 2009) and prior investigations of it (Jorgenson, Miller & Sperber, 1981; Gibbs, 1986), Experiment 3 was designed to test the role that the presence of an echoic-prime plays in irony processing. When a prime was introduced that would later be echoed, an interaction similar to Experiment 1's returned and *despite the presence of decoys*. This supports the Echoic Mention Theory, which assumes that attitude ascription is central to irony comprehension while further defending the idea that it depends on a reference to an available thought.

Are experimental features alone responsible for causing these reported interactions that are thought to rely on ToM? We do not think so since individual differences appear to play a role too. Data from Experiment 1 reveal a correlation between an individual's score on the subscale *social skill* of the AQ and the difference in reading time (between the *Ironic* and *Literal* conditions) of that individual in the second half of Experiment 1. The more social a participant is the faster she is at deciphering ironic statements (when compared to literal readings) in the second half of the Experiment. In other words, the higher one's score is for

the subscale *social skill* of the AQ (indicating greater autistic behavior), the greater remains the difference between ironic readings and literal readings in the second half of Experiment 1. That social abilities appear to be, at least partly, determinative supports our hypothesis that prior mixed results could be due to individual differences. After all, a subject pool in one university could be more asocial than another.

Taken together, these results demonstrate how attitude-ascription plays a modulating role in irony comprehension and in two ways. One, the interaction between *Type of story* and *Part of session* shows that once a reader captures tendencies in these stories, the comprehension of irony is no longer more involving than the comprehension of literal sentences. Second, the social ability of readers correlates with the ability to access the ironic stance.

These data are in line with previous results of our group (Spotorno, Koun, Prado, Van Der Henst, & Noveck, 2012). While using materials that were similar to Experiment 2's, we predicted that ironic readings, when compared to literal readings of the same target sentences, would reveal greater activations in mindreading areas in the brain. This prediction was based on a general consensus that mindreading processes (what the neuroscience literature refers to as *mentalizing*), encapsulate a ToM network that includes the right Temporo-parietal junction (the rTPJ), the left Temporo-parietal junction (the lTPJ), the Precuneus and the Medial Prefrontal Cortex (Saxe, Carey & Kanwisher, 2004; Frith & Frith, 2006; van Overwalle, 2009). While nearly all neuroscience studies rely on interpretations of an actor's actions, Spotorno et al. (2012) showed how ironic readings, too, engage this network. That study, which is -- to our knowledge -- the first neuroscience study to specifically show a role for ToM in language processing adds further credence to our claim here that attitude ascription is central to irony processing.

Despite our optimism that this work is on the right track, there are at least two reasons why we want to be cautious in the interpretation of our data. First, the link between the effect of interaction (*Type of story * Part of session*) and mindreading abilities follows from indirect inferences. We have not included a set of "pure" tests of ToM abilities when looking for a direct correlation between irony processing and ToM and so our ToM-related hypotheses have relied on the (albeit strong) theoretical assumption that ToM plays a critical role in irony processing along with previous behavioral (e.g., Filippova & Astington, 2008) and fMRI (e.g., Spotorno et al. 2012) data. Second, we decided to hold down the length of the experimental session in order to obtain data from non-burdened, vigilant participant and this strategy obliges one to employ a limited number of stories (i.e., 10 for each condition). With these two

caveats in mind, we remain confident with respect to our claims however. We addressed our first concern by introducing the AQ questionnaire and by considering a rich theoretical backdrop that considers ToM central. The second concern is assuaged by the fact that our stimuli have provided us with a reliable and readily interpretable set of data.

That ToM plays a characteristic role in irony processing can explain why previous work has reported that ironic readings are as fast as literal ones. With respect to Gibbs's work (1986), which spearheaded Direct Access arguments, we would argue that its experimental features essentially allow for ToM interventions and perhaps even early in the session. For example, the vignettes testing for Echoic Mention in Gibbs's paradigm came with longer, more detailed stories and were not paired with decoys. These two factors can both encourage a quick habituation to, and ready interventions from, ToM when a reader gets to the target sentences.

We have argued that while linguistic decoding and ToM are critically important to all novel communications, the impact of ToM on comprehension (the mindreading part in reading times) becomes less burdensome over the course of an experiment because earlier experiences make the search for a speaker's intention easier as cases repeat themselves. Once intended readings become more obvious, linguistic decoding essentially becomes a shorthand that allows for a more ready access to ToM; this shorthand minimizes the impact that would follow when the two work together spontaneously. In a comprehension task like ours, the linguistic code expressing irony in predictable contexts can practically become a shorthand to gain access to the speaker's intention.

This "shorthand" hypothesis could arguably be generalized. Consider the *conceptual pacts* literature (see Brennan & Clark, 1996) where interlocutors continue to use a more specific reference from earlier exchanges (e.g., when it paid to distinguish a "red Labrador" from "dog") even though it is no longer directly useful to do so (e.g. when the same red Labrador is the only dog to which to refer in a new context). This persistence is said to exploit *lexical entrainment*, which is the idea that once interlocutors hit upon a common expression to refer to a specific object, they continue to do so even if it appears overinformative later. From our point of view, these data could also be seen as an indication that adopted names for references -- which when coined required relatively deep pragmatic processing -- remain active in later exchanges because they echo the previous uses and at a low cost to ToM. Once downstream in an experiment (or conversation), reference to an earlier spontaneous and effortful pragmatic inference-making step ends up being less effortful than the necessity to carry out new (ToM) processing each and every time in context. To put it

another way, in situations that call for repetition, one can see ToM processes piggy-backing on to the linguistic code in such a way that in the code itself becomes an efficient reference to prior ToM processing. While intriguing, these claims are of course hypothetical and call for experimental justification.

To conclude, we have argued that the interpretation of ironic utterances speeds up when a reader no longer relies predominantly on linguistic decoding (in order to make an attitude ascription) and relies primarily on attitude-ascription instead. The same summary (ironically enough) holds for the psycholinguistic literature. Much work in the psycholinguistic literature has focused on the surface features of ironic statements and these efforts have largely led to a stalemate in explaining the literature's "mixed" results about the speed of irony processing. We argue that once attitude-ascription is taken into consideration, one would expect reductions in latencies and specifically for irony processing. Under specific (and unexceptional) experimental conditions, a reader can eventually rely on ToM to understand an ironic remark to the point that it could be understood as fast as a literal one. However, under more spontaneous conditions (as reflected by the early parts of an experimental session), a single utterance is more likely to engage deeper and longer processing when it is understood ironically rather than literally.

References

Astington, J. W., Pelletier, J., & Homer, B. (2002). Theory of mind and epistemological development: the relation between children's second-order false-belief understanding and their ability to reason about evidence. *New Ideas in Psychology*, *20*(2-3), 131-144.

Banerjee, R. (2000). The development of an understanding of modesty. *British Journal of Developmental Psychology*, *18*(4), 499-517.

Baron-Cohen, S. (1995). *Mindblindness: An essay on autism and theory of mind*. Boston: MIT Press/Bradford Books.

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 …*, *31*(1), 5-17.

Clark, H., & Carlson, T. (1982). Hearers and speech acts. *Language*, *58*(2), 332-373.

Clark, H. H., & Gerrig, R. J. (1984). On the pretense theory of irony. *Journal of experimental psychology. General*, *113*(1), 121-6.

Dews, S., & Winner, E. (1999). Obligatory processing of literal and nonliteral meanings in verbal irony. *Journal of pragmatics*, *31*, 1579-1599.

Filik, R., & Moxey, L. (2010). The on-line processing of written irony. *Cognition*, *116*(3), 421–436. Elsevier.

Gibbs, R. W. (1986). On the psycholinguistics of sarcasm. *Journal of Experimental Psychology: General*, *115*(1), 3-15.

Gibbs, Raymond W. (1994). *The poetics of mind: Figurative thought, language, and understanding*. cambridge: Cambridge University Press.

Giora, R. (1995). On Irony and Negation. *Discourse Processes*, *19*, 239-264.

Giora, R. (1997). Understanding figurative and literal language: The graded salience hypothesis. *Cognitive Linguistics*, *8*(3), 183-206.

Giora, R., & Fein, O. (1999). Irony: Context and salience. *Metaphor and Symbol*, *14*(4), 241–257. Psychology Press.

Giora, R., Fein, O., Kaufman, R., Eisenberg, D., & Erez, S. (2009). Does an "ironic situation" favor an ironic interpretation? In G. Brône & J. Vandaele (Eds.), *Cognitive poetics: Goals, gains and gaps* (pp. 383-399). Berlin/New York: Mouton de Gruyter.

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

Happé, F. G. (1993). Communicative competence and theory of mind in autism: a test of relevance theory. *Cognition*, *48*(2), 101–119.

Happé, F. (1994). An advanced test of theory of mind: Understanding of story characters' thoughts and feelings by able autistic, mentally handicapped, and normal children and adults. *Journal of autism and Developmental disorders*, *24*(2), 129-154.

Howell, D. C. (1997). *Statistical methods for psychology* (4th ed.). Belmont, CA: Duxbury Press.

Ivanko, S. L., & Pexman, P. M. (2003). Context Incongruity and Irony Processing Context Incongruity and Irony Processing. *Discourse Processes*, *35*(3), 241 - 279.

Jorgensen, J., Miller, G., & Sperber, D. (1984). Test of the mention theory of irony. *Journal of Experimental Psychology. General*, *113*(1), 112-120.

Katz, A., Blasko, D., & Kazmerski, V. (2004). Saying what you don't mean social influences on sarcastic language processing. *Current Directions in Psycological Science*, *13*(5), 186-189.

Leslie, A. M. (1987). Pretense and representation: The origins of "theory of mind." *Psychological Review*, *94*(4), 412-426.

Schwoebel, J., Dews, S., Winner, E., & Srinivas, K. (2000). Obligatory processing of the literal meaning of ironic utterances: further evidence. *Metaphor and Symbol*, *15*(1 & 2), 47-61.

Sperber, D., & Wilson, D. (1981). Irony and the use-mention distinction. In P. Cole (Ed.), *Radical pragmatics* (pp. 295-318). New York: Academic Press Elsevier.

Spotorno, N., Koun, E., Prado, J., Van Der Henst, J.-B., & Noveck, I. A. (2012). Neural evidence that utterance-processing entails mentalizing: The case of irony. *NeuroImage*, *63*(1), 25–39. Elsevier B.V.

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: the truth about false belief. *Child development*, *72*(3), 655-84.

Wilson, D. (2009). Irony and metarepresentation. *UCLWPL*, *21*, 183-226.

Wilson, D., & Sperber, D. (2012a). Explaining Irony. In D. Wilson & D. Sperber (Eds.), *Meaning and Relevance*. Cambridge: Cambridge University Press.

Wilson, Deirdre, & Sperber, D. (2012b). *Meaning and Relevance*. Cambridge: Cambridge University Press.

Appendix A: Examples of ironic and literal stories (with the *to-be-echoed-prime* of Experiment 3 in parentheses)

Condition	French (as presented)	English translations
Ironic	Cynthia et Léa chantent dans le même opéra. Le soir de la première, elles se retrouvent au théâtre. Le spectacle commence pile à l'heure. (Avant de commencer Léa dit a Cynthia : « Ce soir notre performance sera fantastique ! ») Durant la représentation, elles font beaucoup de fausses notes. Après le spectacle, Cynthia dit à Léa : « Ce soir on a fait une performance magistrale. » Tandis qu'elles se démaquillent, les deux filles continuent à parler du spectacle. Question : À votre avis, est-ce que la performance est le matin ?	Cynthia and Léa sing together in the same opera. On the night of the premiere they meet at the theatre. The show begins exactly on time. (Before the show starts Léa says to Cynthia: "This night our performance will be fantastic!") During their performance they often sing off key. After the show, Cynthia says to Léa: "Tonight we gave a superb performance." As they take off their make-up they continue to discuss the show. Question: In your opinion, do you think that the performance was in the morning?
Literal	Cynthia et Léa chantent dans le même opéra. Le soir de la première, elles se retrouvent au théâtre. Le spectacle commence pile à l'heure. (Avant de commencer Léa dit a Cynthia : « Ce soir notre performance sera fantastique ! ») La représentation est excellente et les chanteurs sont longuement applaudis. Après le spectacle, Cynthia dit à Léa : « Ce soir on a fait une performance magistrale. » Tandis qu'elles se démaquillent, les deux filles continuent à parler du spectacle. Question : À votre avis, est-ce que la performance est le matin ?	Cynthia and Léa sing together in the same opera. On the night of the premiere they meet at the theatre. The show begins exactly on time. (Before the show starts Léa says to Cynthia: "This night our performance will be fantastic!") The show was excellent and the singers were given a long applause. After the show, Cynthia says to Léa: "Tonight we gave a superb performance." As they take off their make-up they continue to discuss the show. Question: In your opinion, do you think that the performance was in the morning?
Ironic	Lors d'un dîner, Patrick parle à son collègue Pascal de sa fondation contre l'anorexie. Il explique à Pascal sa nouvelle idée. Il lui dit qu'il va lancer une nouvelle campagne de financement. (Patrick lui dit : « La nouvelle campagne sera une grande réussite. ») Quelques mois après ils se retrouvent pour évaluer les résultats décevants de la campagne. Les gens ont donné beaucoup moins cette fois-ci et Pascal dit à Patrick : « Avec cette campagne on a fait un grand coup. » Patrick et Pascal commencent à penser aux nouvelles activités pour la fondation. Question : A votre avis, est-ce que Pascal et Patrick sont collègues ?	While at dinner, Patrick talks to his colleague Pascal about his foundation to combat anorexia. He explains to Patrick his new idea. He tells him that he is going to begin a new fundraising campaign. (Patrick says him: "The new campaign will be a success.") Several months later, they meet again to evaluate the disappointing results of the campaign. People gave much less this time and Pascal says to Patrick: "This campaign has really been a hit." Patrick and Pascal start to think of new activities for the foundation. Question: In your opinion, do you think Pascal and Patrick are colleagues?
Literal	Lors d'un dîner, Patrick parle à son collègue Pascal de sa fondation contre l'anorexie. Il explique à Pascal sa nouvelle idée. Il lui dit qu'il va lancer une nouvelle campagne de financement. (Patrick lui dit : « La nouvelle campagne sera une grande réussite. ») Quelques mois après ils se retrouvent pour évaluer les excellents résultats de la campagne. Les gens ont donné beaucoup plus que les années précédentes et Pascal dit à Patrick : « Avec cette campagne on a fait un grand coup ! » Patrick et Pascal commencent à penser aux nouvelles activités pour la fondation. Question : A votre avis, est-ce que Pascal et Patrick sont collègues ?	While at dinner, Patrick talks to his colleague Pascal about his foundation to combat anorexia. He explains to Patrick his new idea. He tells him that he is going to begin a new fundraising campaign. (Patrick says him: "The new campaign will be a success.") Several months later, they meet again to evaluate the excellent results of the campaign. People gave much more than in previous years and Pascal says to Patrick: "This campaign has really been a hit!" Patrick and Pascal start to think of new activities for the foundation. Question: In your opinion, do you think Pascal and Patrick are colleagues?
Ironic	Luc parle de ses investissements à Alfred qui est courtier en bourse. Luc voudrait investir son argent dans les actions d'une petite	Luc talks about an investment with Alfred, who is a stockbroker. Luc is interested in buying shares of a small company.

	entreprise. Alfred lui explique les avantages et les inconvénients. (Alfred lui dit : « C'est un très bon investissement. ») Luc achète les actions mais un mois plus tard leur valeur a diminué de moitié. Lors de la réunion suivante, Luc en reparle à Alfred et dit : « <u>C'est ce qui s'appelle un investissement rentable.</u> » Pendant qu'ils parlent, les nouveaux cours de la bourse s'affichent sur le portable de Luc. Question : A votre avis, est-ce que Luc veut investir dans une grande entreprise ?	Alfred describes the advantages and inconveniences of such an investment. (Alfred tells him: "This is a really good investment.") Luc buys the stocks, but one month later their value has dropped by half. At the next meeting, Luc talks about it with Alfred again and says: "<u>This is what's called a worthwhile investment.</u>" As they are talking, Luc's laptop displays new stock quotes. Question: In your opinion, does Luc want to invest in a big company?
Literal	Luc parle de ses investissements à Alfred qui est courtier en bourse. Luc voudrait investir son argent dans les actions d'une petite entreprise. Alfred lui explique les avantages et les inconvénients. (Alfred lui dit : « C'est un très bon investissement. ») Un mois plus tard, le cours des actions de Luc a déjà doublé. Au cours d'un déjeuner, il en reparle à Alfred : « <u>C'est ce qui s'appelle un investissement rentable.</u> » Pendant qu'ils parlent, les nouveaux cours de la bourse s'affichent sur le portable de Luc. Question : A votre avis, est-ce que Luc veut investir dans une grande entreprise ?	Luc talks about an investment with Alfred, who is a stockbroker. Luc is interested in buying shares of a small company. Alfred describes the advantages and inconveniences of such an investment. (Alfred tells him: "This is a really good investment.") One month later their value, Luc's stocks have already doubled. Over a lunch, Luc talks about it again with Alfred: "<u>This is what's called a worthwhile investment.</u>" As they are talking, Luc's laptop displays new stock quotes. Question: In your opinion, does Luc want to invest in a big company?
Ironic	Clara et Isabelle doivent décider quel film aller voir au cinéma. Elles remarquent l'affiche d'un film dans la rue. Elles ne le connaissent pas mais décident d'aller le voir. (Isabelle dit : « Ce doit être un beau film. ») Les deux amies achètent les billets et des pop-corn. Le film se révèle être banal et très ennuyeux, Clara dit alors à Isabelle : « <u>Nous sommes allées voir un film formidable.</u> » Elles sortent de la salle et vont s'acheter une glace. Question : A votre avis, est-ce que Clara et Isabelle vont s'acheter une glace ?	Clara and Isabelle must decide which film to see at the cinema. They see a poster for a film outside. They aren't familiar with it but they decide to go see it. (Isabelle says: "This should be a good movie.") The two friends buy tickets and popcorn. The film turns out to be banal and very boring, so Clara says to Isabelle: "<u>We went to see a wonderful film.</u>" They leave the theater and go buy an ice cream. Question: In your opinion, do you think Clara and Isabelle went to buy an ice cream?
Literal	Clara et Isabelle doivent décider quel film aller voir au cinéma. Elles remarquent l'affiche d'un film dans la rue. Elles ne le connaissent pas mais décident d'aller le voir. (Isabelle dit : « Ce doit être un beau film. ») Les deux amies achètent les billets et des pop-corn. Le film se révèle être excitant et surprenant, Clara dit alors à Isabelle : « <u>Nous sommes allées voir un film formidable.</u> » Elles sortent de la salle et vont s'acheter une glace. Question : A votre avis, est-ce que Clara et Isabelle vont s'acheter une glace ?	Clara and Isabelle must decide which film to see at the cinema. They see a poster for a film outside. They aren't familiar with it but they decide to go see it. (Isabelle says: "This should be a good movie.") The two friends buy tickets and popcorn. The film turns out to be exciting and surprising, so Clara says to Isabelle: "<u>We went to see a wonderful film.</u>" They leave the theater and go buy an ice cream. Question: In your opinion, do you think Clara and Isabelle went to buy an ice cream?

Appendix B: Examples of decoys used in Experiments 2 and 3 (with the *to-be-echoed-prime* of Experiment 3 in parentheses)

French (as presented)	English translation
Matéo déménage et doit déplacer un miroir lourd et très fragile. Il demande à Paul de l'aider. Paul est disponible tout de suite. (Paul lui dit : « Pas de problème, le miroir sera en sécurité avec moi ! ») A peine a-t-il soulevé le miroir que ce dernier se brise en mille morceaux. Matéo dit à Paul: « On a fait une grosse bêtise. » Quelques jours plus tard, Matéo fête son emménagement avec des amis. Question: A votre avis, est-ce que Mateo et Paul ont déménagé le miroir sans problème?	Matéo is relocating and has to move a very fragile and heavy mirror. He asks Paul for help. Paul makes himself available immediately. (Paul says: "Don't worry, the mirror will be safe with me!") As soon as Paul lifts the mirror it breaks into a thousand pieces. Mateo says to Paul: "We have made a big mistake." A few days later, Mateo celebrates his move with his friends. Question: In your opinion, do Matéo and Damien move the mirror without problems?
Damien et Myriam vont faire les soldes. Damien n'a pas une idée trop claire sur ce qu'il veut acheter. Myriam lui propose de tester différents magasins. (Myriam lui dit : « On trouvera des habits qui te plairont dans le prochain magasin ! ») En ressortant d'une cabine d'essayage, Damien est vêtu d'une manière très extravagante. En voyant le résultat, il dit à Myriam : « Je n'aime pas ces habits. » Ils décident alors de passer au prochain magasin. Question: A votre avis, est-ce que Damien et Myriam vont chercher des nouveaux vêtements ?	Damien and Myriam go shopping for clothes on sale. Damien doesn't have a very clear idea of what he wants to buy. Myriam suggests that he try different stores. (Myriam tells him: "We will find clothes that you like in the next store.") Damien comes out of one dressing room clothed in a very extravagant manner. Seeing himself in the mirror, he says to Myriam: "I don't like these clothes." They decide to move on to the next store. Question: In your opinion, do you think Damien and Myriam are shopping for new clothes?
Valérie organise une surprise pour l'anniversaire d'une copine. Elle demande à Romain de l'aider. Romain est tout à fait d'accord. (Il dit : « La surprise va être super ! ») Malheureusement, la copine en question découvre le secret une semaine avant la fête. Valérie dit à Romain : « L'effet de surprise est complètement raté. » Cependant la fête a été un succès. Question: A votre avis, est-ce que Valérie cherche à organiser une surprise ?	Valérie organizes a surprise party for a (female) friend's birthday. She asks Romain to help her. Romain agrees. (Romain says: "The surprise will be fantastic!") Unfortunately, the friend in question discovers the secret a week before the party. Valérie says to Romain: "The surprise is completely ruined." Nonetheless the party was a success. Question: In your opinion, do you think Valérie tried to organize a surprise party?
Hugo travaille dans une usine de voiture au service de la sécurité routière. Joël vient le voir le jour des crashs tests. Hugo lui montre les procédures pour les tests. (Joël dit à Hugo : « Les erreurs de construction sont de moins en moins fréquentes. ») Ils réalisent alors un test avec un mannequin assis dans la voiture. Le mannequin est complètement détruit à la fin du test et Hugo dit à Joël : « Il y a sûrement eu une erreur quelque part. » Hugo pense alors aux études faites avec les nouvelles normes de sécurité. Question : A votre avis, est-ce que le mannequin est intact après le test ?	Hugo works at a car factory in the department of road safety. Joel comes to see him on the day of the crash tests. Hugo shows him the testing procedures. (Joël tells to Hugo: "The errors of construction are less and less frequent.") They run a test with a dummy seated in the car. The dummy is completely destroyed at the end of the test and Hugo says to Joel: "There must have been an error somewhere." Then Hugo remembers the studies done using the new safety standards. Question: In your opinion, is the dummy intact after the test?

Appendix C: Examples of positive fillers used in all Experiments (with the *to-be-echoed-prime* of Experiment 3 in parentheses)

French (as presented)	English translation
Jérémy a promis à son fils de lui construire une cabane. Il a acheté du bois de châtaignier pour la fabriquer. Il travailla tout l'après-midi pour la bâtir. (Il dit à son fils : « Cette cabane sera parfaite ! ») Une fois terminée, la cabane est très solide et bien construite. Son fils est très heureux et dit à son père : « Viens jouer avec moi dans la cabane. » Ils jouèrent tous les deux durant tout le week-end dans cette nouvelle cabane. Question : A votre avis, est-ce que la cabane est bien construite ?	Jeremy has promised to his kid to build him a cabin. He bought chestnut wood to build it. He works all the afternoon to finish it. (He says to his kid: "This cabin will be perfect!") In the end, the cabin is solid and well built. His kid is very happy and he tells him: "Come to play with my in the cabin." They play all the weekend long in this new cabin. Question: Do you think that the cabin is well built?
Hubert a invité son collègue Michael qui est un grand amateur de vin. Hubert sort une bonne bouteille. Il demanda à Michael s'il veut goûter le vin en premier. (Il dit à Michael : « C'est une bouteille spéciale. ») Le vin est extrêmement doux, une véritable merveille. Durant le dîner Michael demande à Hubert : « Est-ce que tu es collectionneur de grands vins ? » La soirée fut très agréable pour toutes les personnes présentes à ce dîner. Question : A votre avis, est-ce que Michael boit du vin pour la première fois ?	Hubert has invited his colleague Michael who is a wine expert. Hubert open one of his bottles. He asks to Michael if he wants to taste the wine as first. (He tells to Michael: "This is a great bottle.") The wine is extremely mellow, a real marvel. During the dinner Michael asks to Hubert: "Are you a collector of wine?" The night has been enjoyable for all the people that participate to the dinner. Question: Do you think that Michael drinks wine for the first time in his life?
Pauline a conseillé un restaurant à Sarah et Thomas. Ils viennent à Lyon pour la première fois. Pauline ajoute que le restaurant qu'elle leur propose est bon marche. (Pauline ajoute : « Les plats sont bons et bon marché. ») En effet, les plats sont délicieux et en plus ils ne sont pas chers. Le lendemain, Sarah appelle Pauline et lui dit : « Nous devrions aller ensemble dans ce restaurant. » Les deux filles prirent alors rendez-vous le samedi suivant pour le diner. Question : A votre avis, est-ce que Sarah a apprécié le restaurant ?	Pauline has suggested a restaurant to Sarah and Thomas. They are going to Lyon for the first time. Pauline adds that the restaurant she has suggested them is cheap. (Pauline adds: "The dishes are taste and not expensive.") In fact, the food is delicious and they are not expensive. The day after, Sarah calls Pauline and tells her: "We have to go together to that restaurant." The two girls take an appointment for a dinner Saturday after. Question: Do you think that Sarah likes the restaurant?
Jonathan et Bastien partent à la plage. Bastien prête sa crème solaire à Jonathan. Il a acheté une crème nouvelle juste ce matin. (Il ajoute : « Cette crème est la meilleur ! ») A la fin de la journée, Jonathan n'a pas pris de coup de soleil. En rentrant, Jonathan dit à Bastien : « Ta crème est d'une surprenante efficacité. » Après la plage les deux amis vont manger une pizza dans le centre-ville. Question : A votre avis, est-ce que la crème était inutile?	Jonathan and Bastien are going to the beach. Bastien leads his sunscreen to Jonathan. He has bought a new sunscreen just that morning. (He adds: "This is the best sunscreen ever!") In the end of the day, Jonathan did not sunburn. While they are going back, Jonathan tells to Bastien: "Your sunscreen is really effective." After the beach, the two friends go to eat a pizza downtown. Question: Do you think that the sunscreen has not been useful?

Table 1.

Examples of experimental stimuli from Experiment 1 (translated from French).

Condition	Example
Ironic	Cynthia and Léa sing together in the same opera. On the night of the premiere they meet at the theatre. The show begins exactly on time. During their performance both ladies sing off key. After the show, Cynthia says to Léa: "Tonight we gave a superb performance." As they take off their make-up they continue to discuss the show. Question: Do you think that the performance was in the morning?
Literal	Cynthia and Léa sing together in the same opera. On the night of the premiere they meet at the theatre. The show begins exactly on time. Both ladies sing beautifully and receive a rapturous round of applause. After the show, Cynthia says to Léa: "Tonight we gave a superb performance" As they take off their make-up they continue to discuss the show. Question: Do you think that the performance was in the morning?
Filler	Jeremy has promised to his kid to build him a cabin. He bought chestnut wood to build it. He works all the afternoon to finish it. In the end, the cabin is solid and well built. His kid is very happy and he tells him: "Come to play with my in the cabin." They play all the weekend long in this new cabin. Question: Do you think that the cabin is well built?

Table 2.

Example of a decoy story used in Experiment 2 (translated from French).

Condition	Example
Decoy	Mateo is relocating and has to move a very fragile and heavy mirror. He asks Paul for help. Paul makes himself available immediately. As soon as Paul lifts the mirror it breaks into a thousand pieces. Mateo says to Paul: "We have made a big mistake." A few days later, Mateo celebrates his move with his friends. Question: In your opinion, do Mateo and Damien move the mirror without problems?

Table 3.

Examples of experimental stimuli from Experiment 3 (translated from French).

Condition	Example
Ironic	Cynthia and Léa sing together in the same opera. On the night of the premiere they meet at the theatre. Before the show starts Léa says to Cynthia: "This night our performance will be fantastic!" During their performance both ladies sing off key. After the show, Cynthia says to Léa: "Tonight we gave a superb performance." As they take off their make-up they continue to discuss the show. Question: Do you think that the performance was in the morning?
Literal	Cynthia and Léa sing together in the same opera. On the night of the premiere they meet at the theatre. Before the beginning of the show Léa says to Cynthia: "This night our performance will be fantastic!" Both ladies sing beautifully and receive a rapturous round of applause. After the show, Cynthia says to Léa: "Tonight we gave a superb performance" As they take off their make-up they continue to discuss the show. Question: Do you think that the performance was in the morning?
Decoy	Mateo is relocating and has to move a very fragile and heavy mirror. He asks Paul for help. Paul says: "Don't worry, the mirror will be safe with me!" As soon as Paul lifts the mirror it breaks into a thousand pieces. Mateo says to Paul: "We have made a big mistake." A few days later, Mateo celebrates his move with his friends. Question: In your opinion, do Mateo and Damien move the mirror without problems?
Filler	Jeremy has promised his kid to build him a cabin. He bought chestnut wood to build it. He says to his kid: "This cabin will be perfect!" In the end, the cabin is solid and well built. His kid is very happy and he tells him: "Come to play with my in the cabin." They play all the weekend long in this new cabin. Question: Do you think that the cabin is well built?

Figure legends

Figure 1. Experimental procedure.

Figure 2. Reading times for the Ironic and Literal condition in the first experiment. The axis X reports the distinction between first half (Early) and second half (Late) of the experimental session. The error bars report the standard error. Panel A. Reading times for the critical sentence. The difference in reading times between the Ironic and Literal condition, which is clear in the first half, tends to disappear in the second half. Panel B. Reading times for the wrap-up sentence. The difference in reading times between the Ironic and Literal condition is significant in both the first and the second half of the experimental session.

Figure 3. Correlation between the difference *Ironic – Literal* in the second half of the experiment with the social skill subscale of the AQ ($R = .326$, $p < .05$).

Figure 4. Reading times for the Ironic and Literal condition in the second experiment. The axis X reports the distinction between first half (Early) and second half (Late) of the experimental session. The error bars report the standard error. Panel A. Reading times for the critical sentence. The difference in reading times between the Ironic and Literal condition is significant in both the first and the second half of the experimental session. Panel B. Reading times for the wrap-up sentence. The interaction between the variable *Ironic / Literal* and the variable *Early / Late* is slightly significant ($p = .05$). This means that in the second half of the experiment the difference in reading times between the Ironic and Literal condition is not significant.

Figure 5. Reading times for the Ironic and Literal condition in the third experiment. The axis X reports the distinction between first half (Early) and second half (Late) of the experimental session. The error bars report the standard error. Panel A. Reading times for the critical sentence. The difference in reading times between the Ironic and Literal condition, which is clear in the first half, tends to disappear in the second half. Panel B. Reading times for the wrap-up sentence. The difference in reading times between the Ironic and Literal condition is significant in both the first and the second half of the experimental session.
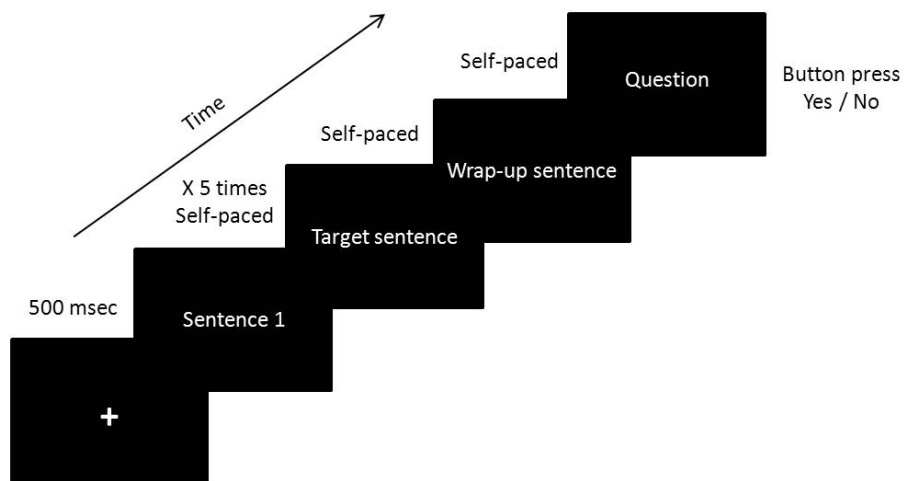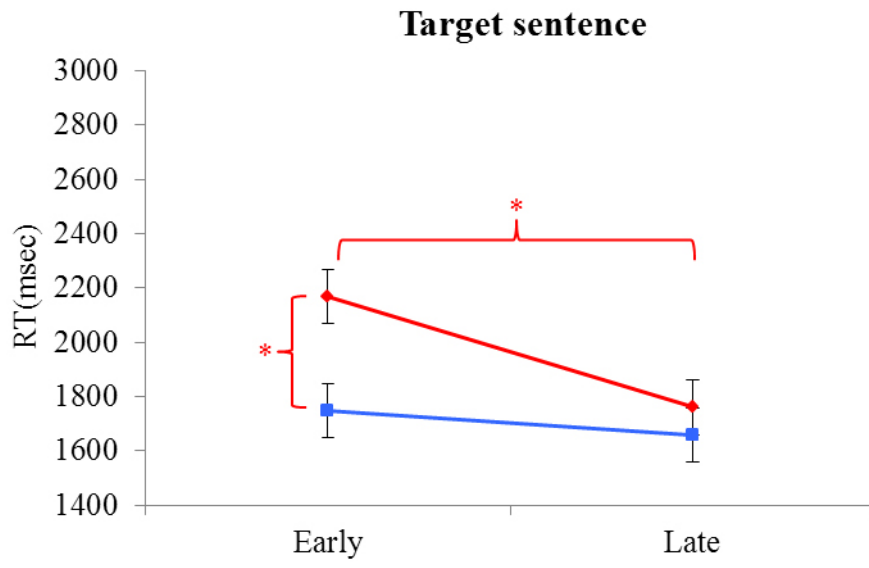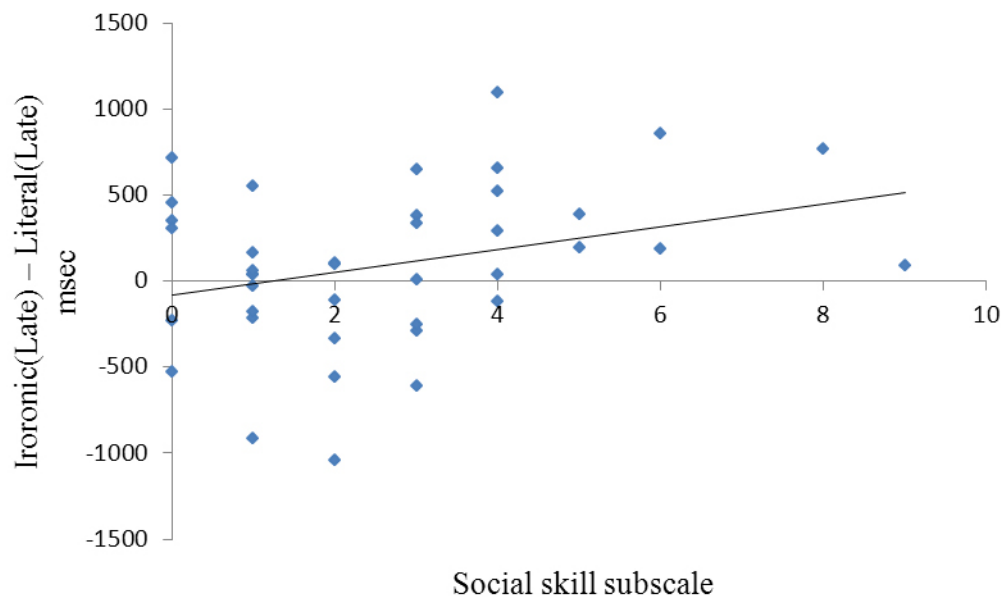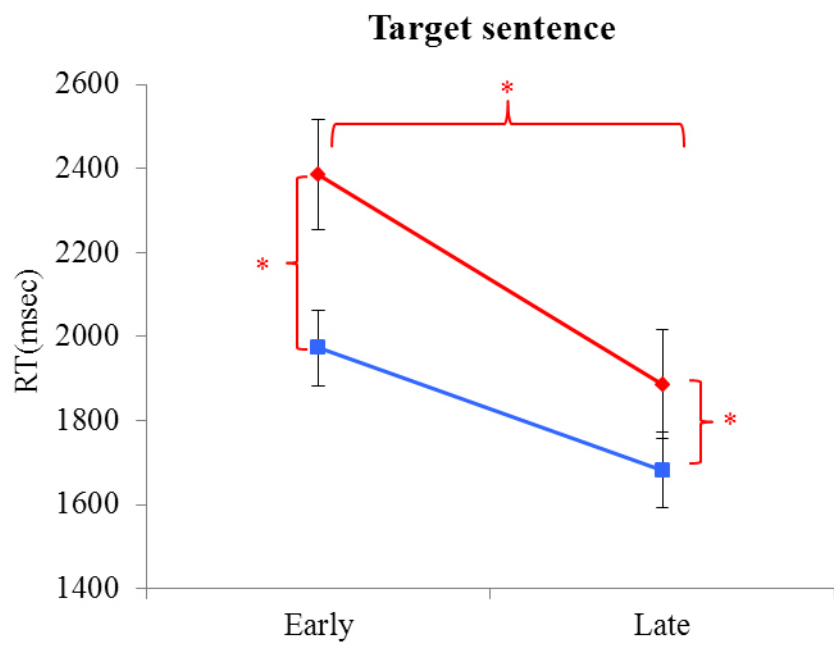
Figure 1.

Figure 2.

Figure 3.
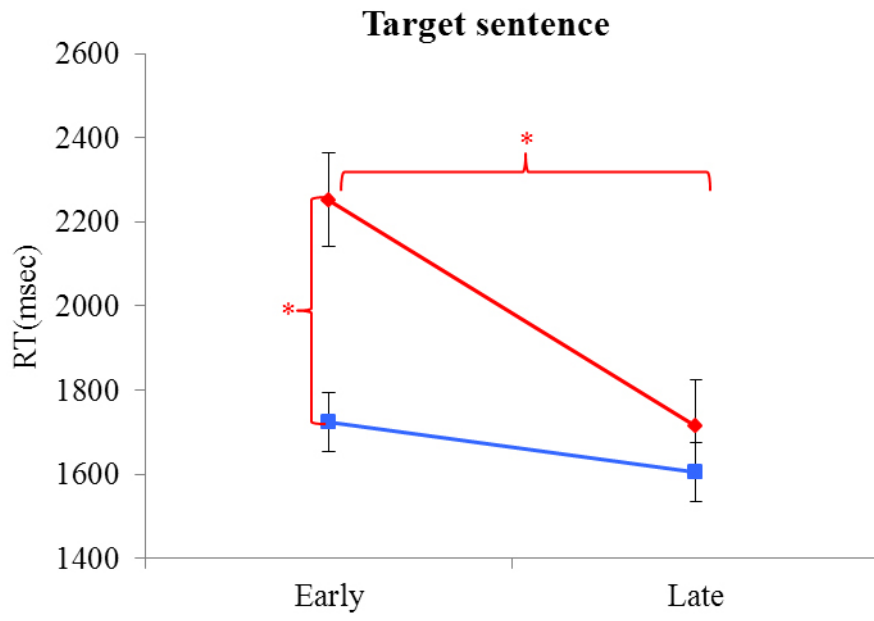
Figure 4.

Target sentence

Figure 5.

IRONY IN OUR BRAIN

5   Irony and the neuro-imaging literature

By looking for a *fil rouge* in the psycholinguistic literature on irony we accomplished our two main goals. First, the three reading time experiments showed that attitude ascription has a role to play in irony processing. Second, we refined paradigm and a set of materials that provide a basis for experimentally investigating irony processing. Now we are ready to turn our attention to the neural basis of the comprehension of ironic remarks. At its most basic neural level, one would expect that an increased demand for ToM in the reading of ironic utterances should be evident through the increased activation of the ToM-related regions. The next chapter briefly reviews some of main features of a rather large literature that concerns the neural basis of ToM.

5.1   The neural basis of Theory of Mind

Based on the work of several researchers (Rebecca Saxe, Simon Baron-Cohen, Uta and Chris Frith, Jason Mitchell and many others), there is now a general consensus that the best candidate regions for the neural underpinnings of the Theory of Mind ability are the right and left temporo-parietal junction (the rTPJ and lTPJ), the medial prefrontal cortex (MPFC) and the precuneus (PC). These four regions are often collectively defined as the "ToM network" and their involvement in ToM-processing has been shown through studies employing different techniques such as fMRI (e.g., Jenkins & Mitchell, 2010; Kampe, Frith, & Frith, 2003; R Saxe & Kanwisher, 2003; R Saxe & Powell, 2006; Rebecca Saxe & Wexler, 2005), transcranial magnetic stimulation (TMS; e.g., Kalbe et al., 2010; Lev-Rana, Shamay-Tsoory, Zangenc, & Levkovitz, 2012) and lesion studies (e.g., Shamay-Tsoory & Aharon-Peretz, 2007; Shamay-Tsoory, Aharon-Peretz, & Levkovitz, 2007; Stone, Baron-Cohen, Calder, Keane, & Young, 2003).

As in the behavioral literature, a large portion of neuroimaging studies on ToM are based on the FBT and so the challenges to the FBT described earlier also affect the investigation of the neural bases of mindreading abilities. The critical question is whether the proposed ToM network is indeed dedicated to Theory of Mind processing; perhaps more general or different processes engage these regions. Considering that the FBT critically requires abilities other than Theory of Mind, such as language and inhibitory control, it is plausible that activations in the ToM network reflect a *mélange* of the three components.

Several studies investigate this issue by employing paradigms that try to separate ToM ability from the others.

5.2   Theory of Mind and language

Support for a link between ToM and language abilities comes from studies on deaf children (e.g., Peterson & Siegal, 1995). It was found that deaf children of hearing parents (i.e., children whose parents are non-native signers) show a delay in passing the false belief task (e.g., Peterson & Siegal, 1999; Wellman & Liu, 2004). In contrast, deaf children of deaf parents (i.e., children whose parents are native signers) do not present this delay (e.g., de Villiers, 2005). Therefore, being exposed to language seems to affect the development of Theory of Mind.

An alternative to assuming that there is a ToM network is to say that activity in these areas reflects some aspect of language processing rather than something related to a mindreading ability. This alternative has been tested by contrasting stories whose comprehension requires the attribution of intentions with stories that do not. Saxe and Kanwisher presented stories requiring the attribution of mental states (but without explicitly mentioning mental states) while the control condition employed stories about mechanical actions that can be interpreted without any reference to beliefs. As predicted, there were greater activations in the ToM regions in the condition requiring the interpretation of mental states as compared to the control condition. Other studies used non-verbal (single-frame or animated) cartoons that contrast conditions in which the attribution of intentions is necessary with conditions in which attributions are not called for (Ciaramidaro et al., 2007; Enrici, Adenzato, Cappa, Bara, & Tettamanti, 2011; Gallagher et al., 2000). All of these studies, which are based on non-verbal material, reported a major engagement of ToM network in the mindreading-condition when compared to the control.
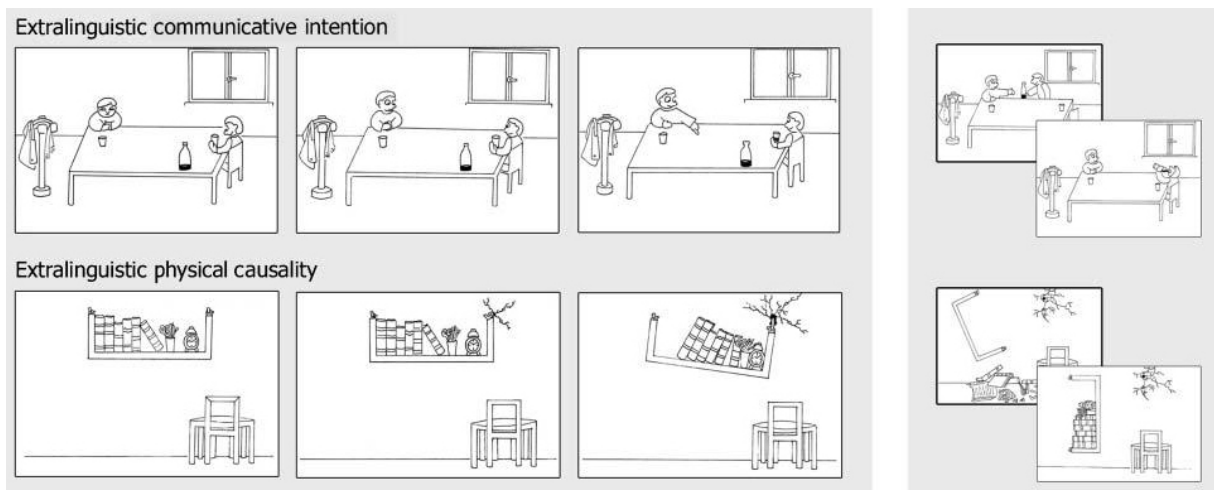
Figure 4. Example of two of the target conditions in the study by Enrici et al. (2011).

There is no evidence that activation in ToM regions directly correlates with the linguistic load of a task, at least in adults. This leads to the hypothesis that while the development of the concepts of desire and belief may need the contribution of language, a mature ToM can work independently from the linguistic system. A critical test for the hypothesis of the autonomy of the ToM network from language is the study of late-acquired aphasia. Typical adults (who presumably have a fully functioning ToM) who then suffer a severe loss of language capacity due to a left hemisphere stroke do not lose ToM (Siegal & Varley, 2006). For example, consider the patient, PH, who became impaired in all aspects of syntax following a massive left hemisphere stroke (Apperly, Samson, Carroll, Hussain, & Humphreys, 2006). He cannot use grammar to understand sentences like "Mary was pushed by Bob" and he also fails to understand embedded sentences. Nevertheless, PH performs perfectly well on non-verbal tests of ToM and even on those that require second-order inferences. Therefore, it seems that people can think about thoughts using non-linguistic strategies. Collectively, these results support the idea that the Theory of Mind network can be independent from the language-related network, at least once the development of mindreading circuits is completed.

5.3 Theory of Mind and executive function

Another critical aspect of the false belief task is that the subject has to contrast two diverging representations of a state of affairs (the actual configuration of the world and the configuration that the character has in mind) and she has to inhibit the representation which is

incoherent with the character's knowledge even if it is the right one. This process strongly engages executive functions and so it is possible that the activation of the ToM network reflects inhibitory/executive abilities more than mindreading ones. One way to investigate this hypothesis is to compare the FBT to another task, the false photographs/maps task, that does not rely on making inferences about intentional agents. The false photograph stories require subjects to represent the (false) content of a physical representation such as a photograph or maps (Zaitchik, 1990); i.e. intentional agents have been replaced by neutral representations. The contrast between the FBT and the

The False-Photographs Task (from Frith, 2001)

false photo/maps tasks has revealed that children with Autism spectrum disorders pass the false photos/maps tasks but not the FBT (e.g., Charman & Baron-Cohen, 1992; Leslie & Thaiss, 1992).

There are good empirical reasons to consider a role for inhibitory and executive functions in the FBT and for the following three reasons. First, as we have already underlined, children younger than 4 years of age are able to pass versions of the task that reduce demands on inhibition (e.g., Lewis & Osborne, 1990; H M Wellman, Cross, & Watson, 2001; Yazdi, German, Defeyter, & Siegal, 2006). Second, individual differences in executive control strongly affect individual children's performance on ToM tasks (e.g., Carlson & Moses, 2001; Sabbagh, Xu, Carlson, Moses, & Lee, 2006). Third, children who fail the FBT also fail tasks that are comparably demanding in terms of executive functions but that do not include any reference to mental states (e.g., Roth & Leslie, 1998; Slaughter, 1998; Zaitchik, 1990).

In order to disentangle ToM from executive processes, researchers such as Rebecca Saxe, Nancy Kanwisher and Josef Perner have conducted fMRI studies in which they contrasted the FBT and false photos/maps tasks, too. The results reveal that the ToM network is more engaged by stories with intentional agents than by stories based on photos or maps (e.g., Perner, Aichhorn, Kronbichler, Staffen, & Ladurner, 2006; R Saxe & Kanwisher, 2003). More generally, the imaging literature on ToM seems to suggest that the rTPJ is the core region of the ToM network (for review see Rebecca Saxe, 2010, but see Mitchell, 2008 for a different view). That said, several studies have found that activity in the rTPJ is also related with attention shift, which is an executive-related skill (e.g., Bledowski, Prvulovic, Goebel, Zanella, & Linden, 2004; Downar, Crawley, Mikulis, & Davis, 2000; Vossel, Weidner, Thiel, & Fink, 2009). More recent investigations have found that the regions that are activated by

thinking about thoughts and the region engaged by attention shift are adjacent but not overlapping (e.g., Scholz, Triantafyllou, Whitfield-Gabrieli, Brown, & Saxe, 2009). The difference is around 10 mm, which is enough to justify the confusion but also enough to suggest a functional difference between the two areas.

Given the above summary, one can say that there are overlaps between ToM-related areas and regions that are engaged by executive function, such as the intra-parietal sulcus and the frontal eye-fields, but these regions are not the brain areas that are labeled as the Theory of Mind network because activations there do not emerge from the selective contrasts that allow one to isolate the regions specifically engaged by mindreading processes (e.g., FBT - false photos and maps task). In the end, one can say that the neuroscientific literature has shown how a distinct set of regions, including the critical rTPJ, emerge as being more sensitive to reasoning about beliefs than to executive control.

5.4   Does a ToM network exist?

Over the last ten years a growing amount of evidence has supported the hypothesis that a specific set of regions is critically involved in ToM processing and so we can safely say that a ToM-brain-network exists. However, it is crucial to specify what we have in mind when we speak about a ToM network. We do not want to claim that the right and left TPJ, the MPFC and the PC are exclusively dedicated to ToM processing. Prior work in the neurobiology of language has led to functional hypotheses about specific aspects of language processing in these regions. For instance, PC activation has been associated with integrating a sentence into context, in line with claims that link activity in the PC with situation model updating (e.g., Speer, Zacks, & Reynolds, 2007). Similarly, MPFC activations are often cited as evidence of general inference making (e.g., by Ferstl & von Cramon, 2002). With the label "ToM-network" we just want to identify a set of regions that is critically engaged in mindreading processing; we are not advocating for a neural-ToM-module.

A different question is what is the specific contribution of each region of the network to ToM processing. Rebecca Saxe proposed that the rTPJ is especially dedicated to thinking about others' thoughts while the MPFC, for example, is engaged in the processing of social information in general (e.g., Saxe, 2006, 2010). To investigate this distinction in one study, for example, Saxe & Kanwisher (2003) presented stories that conveyed socially relevant information that was visible when looking at a person from the outside (like the color of their

hair and clothing) as well as stories that described the contents of another person's thoughts or beliefs. Their idea was that a region that is involved in the processing of all kinds of social information -- and not only reasoning about beliefs -- should be engaged by both kinds of stories. In fact, activity in the MPFC increased with both types of stories (about beliefs and about social information). On the contrary, the rTPJ was activated more by stories about thoughts or beliefs. While Rebecca Saxe proposed that the rTPJ is specifically dedicated to reasoning about thoughts, she argued that the MPFC should be involved in the construction of triadic relations (e.g., Saxe, 2006). Triadic relations link two people and an object of discussion. The ability to recognize and interpret this kind of relation is a crucial part of social cognition, but it is distinct from the ability to interpret others' thoughts.

The investigation of the specific function of the regions involved in ToM processing is still at its earliest stages and the hypothesis that the rTPJ is dedicated to thinking about people's thoughts is still under debate (for a different point of view see, for example, Mitchell, 2008). However, there is now a general consensus that the right and left TPJ, the MPFC and the PC are the best candidates, collectively, for a neural ToM network.

5.5   fMRI studies on irony

As far as we are able to determine, there are seven studies that employ fMRI in order to investigate irony processing on healthy participants (Eviatar & Just, 2006; A. M. Rapp et al., 2010; Shibata, Toyomura, Itoh, & Abe, 2010; Uchiyama et al., 2006; Uchiyama, Saito, Tanabe, & Harada, 2011; Wakusawa et al., 2007; Wang, Lee, Sigman, & Dapretto, 2006). In contrast to the behavioral studies on irony processing, the fMRI literature is not blind to the role that mindreading pays in irony comprehension. Nevertheless, no single study has reported extensive support for the ToM network in the way, say, the false belief paradigm does in Grèzes, Frith, & Passingham, 2004 or in the way that thinking does in the Saxe & Powell (2006) paradigm. An alternative hypothesis to ToM accounts concerns Right Hemisphere (RH) dominance of figurative processing. This second account has been developed on the bases of lesion studies and it has been supported by some fMRI studies on irony processing (e.g., Eviatar & Just, 2006; Wang et al., 2006) but it has been contradicted by others (e.g., Rapp et al, 2010; meta-analyses on figurative language processing by Bohrn, Altmann, & Jacobs, 2012; Alexander M Rapp, Mutschler, & Erb, 2012). In the following, the

thesis will (briefly) explain our position with respect to both the lack of ToM activity in the fMRI studies on irony and the RH hypothesis.

5.5.1 RH hypothesis and figurative language

Several researchers have found that patients with unilateral RH brain damage are significantly impaired in irony comprehension relative to healthy control subjects (e.g., Kaplan, Brownell, Jacobs, & Gardner, 1990; Shamay-Tsoory, Tomer, & Aharon-Peretz, 2005; Tompkins & Mateer, 1985; Winner, Brownell, Happé, Blum, & Pincus, 1998). Giora, Zaidel, & Soroker, 2000 extended this work by including a comparison group of patients with left hemisphere (LH) brain damage and found that patients with RH lesions performed significantly worse on a sarcasm-detection task than those with LH lesions after controlling for the effects of aphasia. Moreover, lesion studies on other forms of figurative language (e.g., Brownell, Simpson, Bihrle, Potter, & Gardner, 1990; Burgess & Chiarello, 1996; Lancker & Kempler, 1987; Winner & Gardner, 1977) had originally been interpreted as an indication in favor of a strong version of the RH hypothesis. In 2005, Beeman (Jung-Beeman, 2005) proposed that the right cerebral hemisphere has a key role to play in processing non-literal expressions. He argued that due to neuro-anatomical differences between the hemispheres, the RH has a general processing advantage for tasks that require the activation and integration of distant semantic concepts, for example, for making long-term predictive inferences, detecting the overriding theme of a text, or understanding metaphors and figurative language. On the other hand, the LH is described as having a predisposition for analytic tasks that require the activation of close, literal semantic associations. In the last two decades a substantial number of papers on figurative language processing tested these hypotheses, but the results are still mixed: several experimental studies have found evidence for the lateralization of figurative language processing (e.g., Faust & Mashal, 2007; Kacinik & Chiarello, 2007; Schmidt, DeBuse, & Seger, 2007; Sotillo et al., 2005; Stringaris et al., 2006). This includes a TMS-study that reported a causal relationship between activity in the right posterior temporal sulcus and the interpretation of novel metaphors (Pobric, Mashal, Faust, & Lavidor, 2008). However, other studies do not find selective RH activations for non-literal language (e.g., Boulenger, Hauk, & Pulvermüller, 2009; Lee & Dapretto, 2006; Mashal, Faust, Hendler, & Jung-Beeman, 2009; A. Rapp, Leube, Erb, Grodd, & Kircher, 2007; A.M. Rapp, Leube, Erb, Grodd, & Kircher, 2004).

Currently, even studies reporting results in favor of a specific RH contribution for figurative language often argue against a strict dichotomy between hemispheres (e.g., Bambini, Gentili, Ricciardi, Bertinetto, & Pietrini, 2011). Two recent meta-analyses on figurative language processing (Rapp et al., 2012 and Bohrn et al., 2012) have considered hypotheses of hemispheric dominance and their results do not support the RH hypothesis. It seems that figurative language processing engages a left lateralized network.

Considering that part of the ToM network, namely the right temporal parietal junction and, partially, the medial prefrontal cortex, are in the right hemisphere, we propose that part of the claims for RH activity in irony processing come from ToM activity, but as part of an extensive, bilateral network. The results of our fMRI have supported our hypothesis (see the paper: "Neural evidence that utterance-processing entails mentalizing: The case of irony").

5.5.2 ToM in the previous fMRI studies on irony

Every fMRI study on irony (through 2011) says -- at least once in its discussion section -- that Theory of Mind probably plays a role in irony processing before saying that their results disconfirm such expectations. While being sympathetic with the hypothesis in this prior papers, we want to determine a) to what extent their own data do not support the claim and; b) what sort of evidence was sought in order to test the idea that ToM is strongly engaged in irony processing.

Before investigating these two issues in greater detail, it is important to be clear about the coordinates of the ToM network. With that in mind, we have employed coordinates from van Overwalle who, in a recent and extensive meta-analysis on neuro-imaging literature on ToM (Van Overwalle, 2009), has outlined the most prototypical coordinates for the ToM-related regions (see the fMRI paper). One can remark that all the previous studies found that the activity in the MPFC increases during irony processing with respect to a control condition, but that is all. Only Shibata et al. (2010) found an increase of activation in the precuneus and none of the previous studies reveal a strong engagement of the (bilateral) temporal parietal junction. Given these pieces of evidence, it is not surprising that prior studies have not found support for claims indicating that the ToM is involved in irony processing.

As the next paper will explain in more detail, the activation of a single region can lead to multiple interpretations and does not allow one to derive strong inferences about one precise claim. In contrast, the activation of an entire previously-described network would

provide more convincing evidence. While it is also true that the co-activation of four brain regions – namely the rTPJ, the lTPJ, the MPFC and the PC -- cannot guarantee that they are engaged in a single process that is necessarily related to ToM, the anticipated involvement of all four regions, instead of just one or two, is arguably strong evidence in favor of ToM-based explanation of irony processing.

Rebecca Saxe (e.g., Saxe, Brett, & Kanwisher, 2006) proposed another method to test whether the "red blobs" from fMRI analysis reveal ToM-related activity or not. Saxe and colleagues developed a short experimental routine based on the contrast between the False Belief Test and the false photograph task that should be added to every experimental session as a sort of little side-experiment. The main idea of this short task is to localize the areas of the brain that are activated by the most prototypical ToM-revealing contrast in the literature for each subject. After that, the coordinates of those areas can be employed as regions of interest (ROIs) and the data of the main task could then be masked with those ROIs. The advantage of this technique is that the ROIs are customized for each subject, providing more precise borders of the ToM-related regions. A specular technique, which still employs the ROIs, is to choose the coordinate of the regions of interest from a meta-analysis on the ToM literature. This strategy loosens the precision of the ToM localizer at the level of the single subject, but it is less sensitive to the limits imposed by the FBT. For example, the meta-analysis by van Overwalle (2009) is based on a large amount of studies, of which only some are FPTs; in this way, the limits of a single paradigm are smoothed over by the others. After all, there are several techniques that can be applied in order to test with more precision if the ToM network is engaged by irony processing. In the next chapter, the thesis describes the strategy that we have employed in our fMRI study.

6. How we try to capture Theory of Mind activity in irony processing

The aim of the following fMRI study is to investigate the extent to which the ToM-network is involved in irony processing. Considering the lack of ToM-related activation in the previous fMRI studies, one should be prevented from expecting a strong engagement of the ToM network. However, in the paper we will show that the lack of ToM-related activation is arguably due to methodological features of the previous studies. In order to avoid the same methodological issues raised about the prior studies, we selected stimuli from the pool of stories we had used in the behavioral paper and we have included decoy stories as we did in the second and third behavioral studies.

We have also taken advantage of the neuro-imaging literature by combining the classic whole brain analysis with the analysis based on regions of interest and the analysis of the functional connectivity between different brain regions. This study represents a critical advance to the thesis because we combined our knowledge about the theoretical accounts, the psycholinguistic literature and the neuro-imaging studies in order to investigate how language and Theory of Mind interact during irony processing.

Neural evidence that utterance-processing entails mentalizing: The case of irony

Nicola Spotorno[a], Eric Koun[a], Jérôme Prado[a], Jean-Baptiste Van Der Henst[a], Ira A. Noveck[a,b]

[a] Laboratory L2C2, CNRS Université Lyon 1 FRE3406, 69675 Bron CEDEX, France

[b] Centre de recherche français à Jérusalem (CRFJ), MAEE-CNRS (UMIFRE 7) B.P. 547, 3

rue Shimshon, Baka, 91004 Jérusalem, Israel

Corresponding author:

Nicola Spotorno, L2C2 CNRS-Université Lyon 1, Institut des Sciences Cognitives, 67 Bd

Pinel - 69675 Bron CEDEX, France.

Phone: +33 (0)4 37 91 12 71

E-mail address: nicola.spotorno@isc.cnrs.fr

Abstract

It is now well established that communicators interpret others' mental states through what has been called "Theory of Mind" (ToM). From a linguistic-pragmatics perspective, this *mentalizing* ability is considered critical because it is assumed that the linguistic code in all utterances underdetermines the speaker's meaning, leaving a vital role for ToM to fill the gap. From a neuroscience perspective, understanding others' intentions has been shown to activate a neural ToM network that includes the right and left temporal parietal junction (rTPJ, lTPJ), the medial prefrontal cortex (MPFC) and the precuneus (PC). Surprisingly, however, there are no studies – to our knowledge – that aim to uncover a direct, on-line link between language processing and ToM through neuroimaging. This is why we focus on verbal irony, an obviously pragmatic phenomenon that compels a listener to detect the speaker's (dissociated, mocking) attitude (Wilson, 2009). In the present fMRI investigation, we compare participants' comprehension of 18 target sentences as contexts make them either ironic or literal. Consider an opera singer who tells her interlocutor: "Tonight we gave a superb performance!" when the performance in question was clearly awful (making the statement ironic) or very good (making the statement literal). We demonstrate that the ToM network becomes active *while* a participant is understanding verbal irony. Moreover, we demonstrate -- through Psychophysiological Interactions (PPI) analyses -- that ToM activity is directly linked with language comprehension processes. The paradigm, its predictions, and the reported results contrast dramatically with those from seven prior fMRI studies on irony.

Keywords: Theory of Mind, Language, Irony, fMRI, PPI, Pragmatics

1. Introduction

Much work in the cognitive sciences has documented our species' ability to interpret other agents' intentions. This mentalizing ability, often called "Theory of Mind" (ToM) (Premack & Woodruff, 1978b), concerns recognizing belief states or attitudes, beliefs about beliefs (second-order ToM) as well as recognizing that one's prior beliefs could be wrong (among other things). In fact, the earliest experimental research on ToM is based on adaptations of the False Belief paradigm taken from developmental psychology in which a participant is required to predict the behavior of a character who has a false belief about the location of a hidden object (Wimmer & Perner, 1983). A central marker of ToM maturity is the ability to recognize that a protagonist's belief-state ought to prompt her to search for the object where she last saw it and not where it actually is (i.e., Sally ought to look for a marble in a basket where she left it even though it had been moved to a box).

Neuroimaging investigations have played an important role in substantiating that Theory of Mind is involved in false beliefs and, more generally, in inference-making about other agents. Through a variety of tasks, there is now general agreement that one can point to a neural ToM network that includes the right temporal-parietal junction (rTPJ), the left temporal-parietal junction (lTPJ), the medial prefrontal cortex (MPFC) and the precuneus (PC) (for reviews, see Mitchell, 2009; Saxe, Carey, & Kanwisher, 2004; Van Overwalle & Baetens, 2009; also see Frith and Frith, 2006 who would add the temporal poles as well). According to Saxe (2006), the right TPJ and the MPFC are especially central to ToM processing. The former is thought to be selectively recruited for reasoning about, and the interpretation of, the content of mental states (how the state of affairs is represented, i.e. what the person believes to be true of it). The latter (especially the dorsal part of the MPFC) is considered to be implicated in the representation of triadic relations, which can be viewed as *You, Me and a state of affairs* and which allows one to correctly interpret situations where two people are talking, thinking or working on a shared third "object."

Neurologically speaking, there are three ways in which research has investigated links between language and ToM. One is through studies that investigate the extent to which language and ToM abilities co-develop during the life span by determining whether a deficit in one affects the other (e.g., Malle, 2002). A second way is to investigate the link between pragmatic deficits and Autism Spectrum Disorders. Those on the Autism Spectrum are thought to be less capable at accomplishing linguistic tasks that depend on ToM (see Happé,

1993; for a review see Tager-Flusberg, 2000; though see (Chevallier et al., 2012, 2011; Chevallier & Wilson, 2010). Finally, others have considered the connection between ToM-related regions and the -- intended and unintended – communication of attitudes. For example, Frith and Frith (2006) reviewed several studies that show how ToM areas are activated when a participant is essentially asked to make inferences about others. For example, the Friths cite a study from Mitchell, Macrae, & Banaji (2006) in which the authors asked participants to predict attitudes of two individuals, one of whom was described as liberal and the other as conservative (e.g., ''would he enjoy having a roommate from a different country''). Those sharing the same outlook prompt a pattern of results that differ from those who do not share the same leanings. This kind of work has ultimately been concerned with the segregation of activity associated with mentalizing and specifically with respect to the MPFC, one of the main regions of ToM activity.

The present work also aims to establish a neural relationship between ToM and language but more specifically by investigating how ToM is engaged *while* participants process utterances. No prior work -- as far as we can tell -- has aimed to establish such a link experimentally through neuroimaging (though see Saxe, 2009, for some speculations about such a potential link). Given the extensive findings on ToM, it would be eminently sensible to assume that one would find neurological evidence linking linguistic triggers and ToM processing. After all, an utterance is typically a starting point for understanding an agent's intention and ToM is crucial for filling the gaps between *what is said* and *what is meant* (Carston, 2002; Sperber & Wilson, 1986). We attempt to establish this link by investigating verbal irony, which is a figure of speech that most obviously underlines how one need exploit ToM in utterance comprehension.

In what follows, we describe properties of irony as well as explain why its processing would make for an ideal linguistic-pragmatic object for investigating the ToM network. We briefly review the existing neuroimaging literature on irony and show how there is little support so far for the notion that irony provokes activity in a ToM network. We argue that this negative result is due largely to the methods and materials in the prior experiments. We then briefly summarize how the cognitive neuroscience literature currently views irony processing before turning to our experiment, which aims to establish that ironic utterances specifically prompt ToM activity.

1.1 <u>Processing irony</u>

Ironic utterances provide prima facie evidence that a listener is required to, not only go beyond the literal meaning of an utterance but to, convey the speaker's attitude. That is, irony prompts a listener to understand the speaker's mental state about the proposition expressed. It is not surprising then that theorists generally highlight how irony comes with some form of attitude ascription (Grice, 1989; Clark & Gerrig, 1984; Wilson, 2009). Grice, for one, suggested that irony involves the expression of a "hostile or derogatory judgment or a feeling such as indignation or contempt" (Grice, 1989: 53). For Wilson (2006), the dissociativeness of irony is central to the echoic-mention theory (Wilson & Sperber, 1992) of irony, as she explains in the example of Mary who, after a difficult meeting, says "That went well":

> Mary might use ["That went well"] to communicate that it was ridiculous of her
> to think that the meeting would go well, stupid of her friends to assure her that it
> would go well, naïve of her to believe their assurances, and so on. Mary *echoes* a
> thought or utterance with a similar content to the one expressed in her utterance,
> in order to express a critical or mocking attitude to it.

It is clear -- intuitively and theoretically -- that the understanding of irony requires that one access a speaker's intention as she expresses a dissociative attitude.

There are currently seven studies that employ functional neuroimaging techniques to investigate irony processing on healthy participants (Eviatar & Just, 2006; Rapp et al., 2010; Shibata, Toyomura, Itoh, & Abe, 2010; Uchiyama et al., 2011, 2006; Wakusawa et al., 2007; Wang, Lee, Sigman, & Dapretto, 2006; see Table 1 for a description and summary of the prior neuroimaging studies). Remarkably, no single study reports extensive activity in the ToM network in the way that neural ToM tasks do (cf. Saxe & Powell, 2006). Instead, one finds either a) no overlap with ToM regions or b) only partial overlap.

As an example of the first sort, consider one recent study (Uchiyama et al., 2011). The authors developed vignettes that had ironic comments as endings, e.g. see the one in (2a) below, and they also worked out slightly different contexts that would render such comments more literal, e.g. see the one in (2b):

(2a) The woman was not a good cook and was taking up to an
        hour just preparing the ingredients. Her mother-in-law, who

was watching how she was doing, said to her: "You're very
skillful."

(2b) The woman was a good cook and was preparing dinner
efficiently. Her mother-in-law, who was watching how she
was doing, said to her: "You're very skillful."

Utterances such as those at the end of (2a) activated participants' subcortical and limbic regions when compared to the non-ironic versions of utterances such as those in (2b). These regions have no overlap with those considered to be part of the ToM network.

As an example of the second sort, consider Eviatar and Just (2006) who compared participants' reading of metaphoric, ironic and literal utterances when they were presented as the final passage of 4-sentence-long vignettes. They found limited extra activation in the regions near the Right TPJ (coordinates x = 51 y = -26 z = 5) in the *Ironic* condition, but not much else (also see Wang et al., 2006). In short, findings from prior studies do not reveal extensive overlap with the ToM system.

Insert Table 1 about here

In the absence of ToM activity in these studies, one is of course prevented from claiming that ToM is involved in irony processing, let alone language. It is our view, however, that the prior investigations have not reported extensive ToM activity because the methods and materials were not ideal for discovering it. Below, we summarize three general features of these studies and consider how they have arguably prevented the literature from determining that verbal irony prompts neurological ToM activity.

First, the presentation of vignettes in the prior fMRI studies has not been optimal for the two following reasons. One is that vignettes and their target utterances are almost invariably short, ranging from two to at most four sentences (Eviatar & Just, 2006; A. M. Rapp et al., 2010; Shibata et al., 2010; Uchiyama et al., 2006, 2011; Wakusawa et al., 2007; Wang et al., 2006). Arguably, such brevity gives a participant a limited amount of time to appreciate the background of a story and its eventual irony (as in 2a above). These stimuli were no doubt developed in order to fit into the parameters of a typical fMRI session. However, judging from prior psycholinguistic investigations, an ironic statement in a laboratory setting requires

contextual development and typically more background. For example, Gibbs (1986) presented stories that were up to seven sentences long and even recent EEG experiments (whose experimental conditions require 30 trials or more) include longer vignettes (see Regel, Gunter, & Friederici, 2011). The differences across the neurological and behavioral literatures do not end there, of course. The other drawback of the neuroimaging studies is that the uptake of the vignettes is typically outside the participant's control (a vignette is presented as a block or else at a speed pre-determined by the experiment). This is unlike the tasks in the behavioral literature, which are usually self-paced. Such features risk taking away from the naturalness of reading texts by constraining a participant's ability to make inferences on-line.

Second, ironic items in the neuroimaging studies predominate the stimuli in two ways. One way is that ironic items are practically telegraphed in the context of these experiments. That is, it appears that ironic utterances in existing neuroimaging studies are systematically cued by negative events while literal uses of similar utterances are not. The upshot is that any effort to mentalize, the very activity that is being investigated, risks becoming short-circuited over the course of an experimental session. The other way is that ironic materials are highly prominent in most of these fMRI investigations. Whereas frequency estimates indicate that irony represents 8% of conversational turns in talk among friends (Gibbs, 2000) and that readers of contemporary American literature can encounter, on average, an irony every four pages (Kreuz, Roberts, Johnson, & Bertus, 1996), the proportion of ironies over the course of a typical fMRI study is much higher. For example, ironic targets represent 1/3 of the stimuli in the studies of Eviatar and Just (2006) and Shibata and colleagues (Shibata et al., 2010). To some extent, these issues can be dealt with extra filler items but rarely are.

Third, the studies do not systematically take advantage of the fact that studies on irony come with an ideal control in which minor modifications to the context can allow one to use the very same sentence as an ironic remark or as a non-ironic, literal one. In two studies (e.g., Wakusawa et al., 2007; Shibata al., 2010), ironic versus literal stories are not designed from common contexts. In three other studies (Eviatar & Just, 2006; Uchiyama et al., 2006; Wang et al., 2006), as Rapp and colleagues (Rapp et al., 2010) point out, *Ironic* and *Literal* conditions are not directly compared. For example, Uchiyama and colleagues' (Uchiyama et al., 2006) *sarcasm-detection* measure was determined by contrasting, on the one hand, the sarcastic and the non-sarcastic remarks together and, on the other, a control sentence that was "unconnected." This does not allow one to isolate the activity linked exclusively to the sarcastic remark.

It is not surprising then that current neurological accounts of irony do not emphasize ToM processing. Instead, the literature underlines more general processes about figurative language that coincidentally involve the Right Hemisphere (RH). For example, Gernsbacher and Robertson (2004) attribute "narrative construction" to the RH while Long & Baynes (2002) use investigations of impairments to claim that the RH is involved in "discourse representation" (for a different perspective, cf. Bambini, Gentili, Ricciardi, Bertinetto, & Pietrini, 2011; Rapp, Leube, Erb, Grodd, & Kircher, 2007). Although the comprehension of irony seems to be especially affected by lesions in the right hemisphere (e.g., Brownell, Carroll, Rehak, & Wingfield, 1992; Shamay-Tsoory, Tomer, & Aharon-Peretz, 2005), the precise role that the RH plays in these impairments remains largely descriptive.[4] Of course, if one *were* to find extensive ToM activity, it would not be inconsistent with accounts that emphasize the right hemisphere's role in figurative language generally; after all, proposed ToM networks include the crucial rTPJ. Our ToM account would provide some precision to claims about RH activation.

In the current investigation, our main question is the following: To what extent does irony comprehension recruit the bilateral ToM network? According to our hypothesis, when a target utterance is part of an *Ironic* condition as opposed to a *Literal* one, one ought to find evidence that covers the entire ToM network (the rTPJ, the lTPJ, the MPFC and the PC). Negative evidence, of course, would imply that the prior studies were on the right track or that perhaps ToM is secondary or even irrelevant to irony processing.

In an effort to find connections between ToM regions and language processing we anticipate using another investigatory technique, the psycho-physiological interactions analysis (PPI) which is designed to determine whether the functional connectivity of an *a priori* determined region interacts with other brain areas as a function of an experimental condition (Friston et al., 1997). In the current study, we aim to investigate the relationship between ToM areas and the left IFG, which is implicated (non-controversially) in language comprehension (e.g., see Hagoort, 2005). More specifically, if activity of ToM regions and language regions co-vary as a function of the *Ironic* v. *Literal* conditions, then it would imply that ToM activity is directly linked to language use.

---

[4] One current debate pits at least two positions against each other. One is that RH patients have deficits in making counterfactual inferences, which are often necessary for understanding speakers. The other is that RH patients have specific difficulties in making inferences about a speaker's mental states (see McDonald, 1999 for a review).

To summarize, our study is designed to investigate the interaction between language processing and the ToM network and by using irony as a test bed. We designed the protocol in the most rigorous and ecological way possible given the experimental constraints imposed by fMRI methodology. We prepared our study so that we could address each of the criticisms raised about the prior studies and while using techniques that could determine the extent to which ToM is implicated in language through irony.

2. Materials and methods

2.1 Participants

Twenty healthy participants, who were students from the University of Lyon, participated in the study (12 females and 8 males). All participants (whose mean age was 22) were native French speakers, were right-handed and reported to have normal vision and no history of mental illness. Our protocol was accepted by the local ethics committee and each participant passed a medical visit and gave informed consent prior to the experiment.

2.2 Materials

Twenty story-frameworks were created (in French) that led to a target sentence that could be interpreted either as ironic or as literal as a function of a minor modification made to the prior context. In the *Ironic* condition, a target sentence (e.g., "Tonight, we gave a superb performance.") was preceded by a negative context (e.g., a terrible performance) whereas in the *Literal* condition the target sentence was preceded by a positive context (e.g., an impressive performance). Otherwise, the introductory sentences and the wrap-up sentences of any given framework was the same for both conditions. Each story included the following six features.

First, all stories were seven lines long, each having a maximum length of 91 characters (spaces included) in order to fit into one line on a screen. Second, the stories described an everyday situation and an exchange between two characters who know each other casually (i.e., we avoided situations that presumed close relationships among interlocutors). Third, the first three sentences introduced the two characters and the situation. Fourth, the fourth and fifth sentences described the development of the situation that can be either positive (in the literal version) or negative (in the ironic version). These were the only two lines that could

potentially change with respect to condition. Changes were designed to be as minimal as possible while keeping the stories sensible. Fifth, the sixth line was designed to be the target sentence. The length of all target sentences was between 10 and 12 syllables whose number of words ranged from 6 to 10, with a mean of 7.4. Crucially, the target sentence (line 6) is exactly the same in both *Ironic* and *Literal* conditions. Finally, the seventh line was an ordinary wrap-up conclusion of the story that makes sense for both the *Ironic* and *Literal* conditions.

As we indicated, we aimed to block a link between negative contexts and ironic remarks through what we refer to as *decoy* stories. The structure of the six decoys was the same as the ironic stories (7 sentences in which a negative event occurs), except that the target sentence was banal. For example, the decoy story in Table 2 describes how one character drops a mirror, which leads the other character to remark "We have made a big mistake." Like in the *Ironic* and *Literal* conditions, the target sentence in the decoy stories is between 10 and 12 syllables.

Each participant read 10 ironic stories, 8 literal stories and 6 decoys. For each participant, the 18 critical (non-decoy) stimuli were extracted randomly from a pool of 20 frameworks that could each be the basis of either an ironic or literal target sentence. The 6 decoys remained the same for each participant. There were also a further 36 filler items (which also consisted of 7-line long stories about everyday situations as well; these were part of another experiment on text comprehension).

These stories were drawn from a corpus that has been developed and tested repeatedly as part of a cognitive study of irony. To verify that the specific stimuli used here were perceived as intended, a rating study was conducted on the 46 stories (2 from each of the frameworks plus the six decoys) with 26 participants (13 women) whose ages ranged from 19 to 35 (with a mean of 27) and who did not participate in the imaging study. Whereas the 40 *Ironic* and *Literal* stories were pseudorandomized and balanced across two lists, the 6 decoys were included for each participant. Participants were asked to read each story and rate the extent to which the target sentence was ironic on a scale from 1 (not at all ironic) to 5 (very ironic). Ironic target sentences were rated as highly ironic (mean of 4.5), while literal sentences and the banal lines from the decoy stories were rated as low on the ironic scale (1.2 and 1.4, respectively). Repeated measure ANOVAs showed significant differences between (i) the *Ironic* and *Literal* conditions and (ii) the *Ironic* condition and *Decoys* (both at $p < .001$,

corrected for multiple comparisons using the Tukey method). The comparison between the *Literal* condition and the *Decoys* was not significant ($p = .1$)

A yes/no comprehension question followed each item (regardless of whether it was a critical or filler item). The question was about some detail in the story that made no reference to the target sentence whose goal was to ensure that the participants were paying attention to the stories. For half of the questions the correct answer was "yes" and for the other half the correct answer was "no" (see Table 2 for an example of all conditions and questions and the Appendix for further examples).

Insert Table 2 about here

2.3 Procedure

Stimuli were prepared with Presentation 11.0 software (Neurobehavioral Systems, www.neurobs.com) and projected onto a translucent screen with a Canon Xeed SX50 projector. The screen was viewed through a mirror. Participants performed the experiment in four runs of 15 stories each. The maximum duration of each run was of 12 min. Each trial started with the presentation of a visual fixation mark (a central cross) in the center of the screen. The fixation mark was red for 7 seconds, orange for 1 second and then green for 1 sec (see Figure 1). The participant read the stories line by line (i.e., sentence by sentence) in a self-paced manner (i.e., each sentence remained on the screen until the participant pressed a key). The interval between the disappearance of a sentence and the presentation of the next one was 500 msec. After the last sentence (line 7) disappeared, the central cross (this time in white) reappeared for 500 msec. The question was then presented and the participants pressed one of two buttons on a keypad (yes/no response). Variable periods of visual fixations (between 2000 and 4000 msec) were added at the end of each trial to introduce jittering. The presentation order of the stories was pseudo-randomized. This means that the number of ironic and literal stories, decoys and fillers was balanced among the sessions. Two ironic stories, 2 literal stories and 2 decoys were presented in runs 1 and 3; three ironic stories, 2 literal stories and 1 decoy were presented in runs 2 and 4. Each stimulus was displayed in a left-justified manner at the center of the rear projection screen. Participants were instructed to read at a normal rate and to respond as accurately as possible to the questions. The

experimental session began with 3 training trials, which do not include ironies. All told, a typical session lasted a little less than an hour (including breaks).

Insert Figure 1 about here

2.4 Imaging Procedures

Images were collected using the 1.5T MRI system (Siemens Sonata Maestro Class; Siemens, Erlangen, Germany) of the CERMEP Imagerie du vivant in Lyon. The fMRI blood oxygenation level dependent (BOLD) signal was measured using a T2*-weighted echo-planar sequence (repetition time [TR] = 2500 msec, flip angle = 90°, echo time [TE] = 60 msec). Twenty-six axial slices (4.40-mm thickness, field of view = 23 cm, 64 64 matrix) were acquired per volume. Following functional image acquisition, a high-resolution T1-weighted anatomical image (TR = 1880 ms, TE = 3.93 ms, FOV = 256 mm, flip angle = 158, 176 x 256 x 256 matrix, slice thickness = 1 mm) was collected for each participant.

2.5 Analysis

2.5.1 Behavioral data

The answers to the questions were analyzed in order to ensure that participants were paying attention to the stories. Each participant answered the comprehension questions correctly at rates that were higher than those predicted by chance (which is a 70% hit rate). The mean rate of correct responses overall was 91.5% (the lowest being 85%). Therefore, all participants were included in the analyses. Statistical analyses were made on the reading times for the target sentences (line 6), which were log transformed. Reading times that were 2.5 standard deviations above or below the mean were considered outliers. Given these constraints, we rejected approximately 4.5% of the data from the behavioral and fMRI analyses.

2.5.2 fMRI Data

fMRI data were analyzed using SPM8 software (Wellcome Department of Cognitive Neurology, London, UK, www.fil.ion.ucl.ac.uk). Each run contained 124 functional volumes after rejecting the first four scans to eliminate nonequilibrium effects of magnetization. Functional images were corrected for slice acquisition delays and were spatially realigned to the first image of the first session on a voxel-by-voxel basis so as to correct for head

movements. The realigned functional images and the anatomical scans for each participant were then normalized into a standard stereotaxic space by using the Montreal Neurological Institute (MNI) template. The functional images were spatially smoothed with an isotropic Gaussian filter (8-mm full width at half maximum). The event-related statistical analysis was performed according to the general linear model (Josephs, Turner, & Friston, 1997) using the standard hemodynamic response function provided by SPM8. Events were time-locked to the appearance of the target sentence (sixth line of the stories). The other sentences of the story were modeled together. The time series data were high-pass filtered (1/128 Hz) and serial correlations were corrected by an autoregressive AR (1) model. Each activation event was categorized according to the experimental variables. Random effects analyses were applied to individual contrasts to account for between-participants variance and to generalize to the population as a whole. The activations reported survived a voxel-level threshold of $p < .001$, uncorrected for multiple comparisons, and a cluster-level threshold of $p < .05$, corrected for multiple comparisons using the FDR method. The SPM8 coordinates were converted from MNI coordinate space into Talairach space (www.mrc-cbu.cam.ac.uk/Imaging/Common/mnispace.shtml) and localized using the Talairach atlas (Talairach & Tournoux, 1988).

Four regions of interest (ROIs) were defined based on two meta-analyses of the ToM network (Van Overwalle 2009; Van Overwalle & Baetens, 2009). More specifically, four sets of coordinates were taken from Figure 2 of Van Overwalle and Baetens (2009), in which the authors report the coordinates that are prototypically considered to be part of the ToM network according to the meta-analysis from Van Overwalle (2009). The ROIs were spheres of 6 mm radius centered in: x = 50 y = -55 z= 25 (rTPJ), x = -50 y = -55 z = 25 (lTPJ), x = 0 y = -60 z = 40 (PC) and x = 0 y = 50 z = 20 (MPFC). Mean percent signal change (PSC) was extracted for each participant and condition using the SPM toolbox Marsbar (http://marsbar.sourceforge.net/). Specifically, mean activity in these regions was defined as the average amount of fMRI activity from 2.5 sec to 7.5 sec following the target sentence as measured with a Finite Impulse Response (FIR) model.

2.5.3 Psychophysiological interaction (PPI) analysis

A PPI analysis was applied to isolate brain areas (targets) showing an activity that can be explained in terms of an interaction between the influence of a distal area (seed) and an experimental parameter (Friston et al., 1997). The goal was to investigate the influence that *a*

priori ToM-related seed regions could exert over other target brain areas in relation to the contrast between the *Ironic* and the *Literal* conditions (i.e., a measure of effective connectivity). For a PPI analysis to be optimal, however, it is important to ensure that the activity in the seed region is not correlated with the contrast of interest. In our case, none of the seed regions that would be used in the PPI analyses should be activated in the *Ironic >* *Literal* contrast. Because we observed greater activity for *Ironic* than *Literal* stories in all of the ROIs defined above (see **Results**), these regions could not be used in the PPI analyses as seeds. To define other seed regions that are still in the relevant ROI's (but potentially inactive with respect to the *Ironic > Literal* contrasts), a two-step procedure was adopted. First, for each ToM region, we obtained alternative coordinates by averaging (by hand) the peak coordinates reported in Table 1, section 11 in Van Overwalle & Baetens (2009). The resulting average coordinates were x = 0 y = 55 z = 6 (ventral part of the MPFC), x = -51 y = -60 z = 26 (left TPJ), x = 54 y = -49 z = 22 (right TPJ) and x = -1 y = -56 z = 33 (PC). Second, we tested whether a significant difference between *Ironic* and *Literal* stories was observed at each of these peaks and only performed PPI analyses with the regions that were not associated with differential activity. Although none of these peaks were present in the whole-brain contrast of *Ironic* vs. *Literal* stories, simple t-tests revealed that *Ironic* stories were associated with more activity than *Literal* stories in both the left and right TPJ (Left TPJ: t(19) = 2.30, *p* = .04; Right TPJ: t(19) = 2.45, *p* = .024). However, no difference was observed in the MPFC (t (19) = 1.18, *p* = .25) and PC (t (19) = 1.79, *p* = .09). Therefore, we only conducted PPI analyses with the MPFC and PC as seed regions. Note that the lack of difference between *Ironic* and *Literal* stories in these particular regions highlights the fact that the regions identified in Overwalle first meta-analysis of ToM are rather large and may not be homogeneously more active during *Ironic* than *Literal* stories.

The above approach amounted to extracting the first eigenvariate time series from 6 mm radius spheres located in the ventral part of the MPFC (center of mass: x = 0 y = 55 z = 6) and the PC (center of mass: x = -1 y = -56 z = 33). Each regional time series served as the first regressor in a distinct PPI analysis (i.e., the "physiological" part of the PPI). Next, we created a second regressor indicating whether each story was *Ironic* or *Literal* (the "psychological" parts of the PPI). Lastly, we created a third regressor reflecting the interaction between the physiological and psychological factors (i.e., the "interaction" parts of the PPI). To compute this interaction term, we first deconvolved the BOLD signal in the seed region by using a Bayesian estimation algorithm (Gitelman, Penny, Ashburner, & Friston, 2003). We then

multiplied the story type and deconvolved seed activity regressors to produce the interaction term. This interaction term was then convolved with a standard HRF. The effect of the interaction term was investigated for each participant and entered into a standard random effect group analysis at the second level (PPI maps were thresholded at an uncorrected voxel-level threshold of $p < .005$, and at a cluster level threshold of $p < .05$ corrected for multiple comparisons using the FDR method). We then determined which of the regions identified in the PPI maps overlapped with the clusters found in the activity analysis for the same contrast by applying a mask that includes all the regions that were activated in the whole brain analysis for the contrast *Ironic > Literal* on to the results of the PPI analysis.

3. Results

3.1 Reading times: Irony vs. Literal

A repeated measures ANOVA performed on the reading times of the target line (line 6) showed that participants took more time reading the target sentences in the *Ironic* condition (2373 msec) when compared to those in the *Literal* condition (2119 msec), $F1 (1,19) = 14.43$ $p < .01$; $F2 (1,17) = 6.6$ $p < .05$, where F1 refers to an analysis by participants and F2 by stories. This is in keeping with the reading time measures in ongoing, behavioral self-paced reading studies in our laboratory.

3.2 fMRI: Irony vs Literal

The *Ironic > Literal* contrast showed greater activity in both the posterior and ventral parts of the MPFC (from the most dorsal Brodmann area 6 to the more ventral 9), the bilateral IFG (especially in its triangular and orbitalis parts; Brodmann areas: 45, 46, 47), the left Insula, the bilateral TPJ (Brodmann areas: 40), the right DLPFC (Brodmann areas: 8) and the right middle temporal gyrus (Brodmann area: 21). See Table 3 for a summary and Figure 2 for brain images. Middle and posterior cingulate cortex and bilateral PC were also more activated in the *Ironic* than in the *Literal* condition, but only at a voxel-level threshold of $p < .005$ (uncorrected for multiple comparisons) and a cluster-level threshold of $p < .05$ (corrected for multiple comparisons). The *Literal > Ironic* contrast did not show any significant activations.

We also analyzed the contrasts (on the sixth line) of *Ironic* vs. *Decoy* stories and *Literal* vs. *Decoy* stories. These contrasts were not associated with any activated brain regions. Decoy stories, however, were only included in the design to block a link between negative contexts and ironic remarks and were not intended to be analyzed in fMRI contrasts. Moreover their lengths and meanings were not designed to be comparable in any way with the Ironic and Literal stories. Their purpose was to keep participants from anticipating an ironic response (to maintain novelty with respect to ironic remarks). This is why only 6 decoy stories were included in the design of the experiment to start with. The lack of difference between the experimental stories and the decoys might thus be due to insufficient statistical power. These contrasts, however, were not central to the present experiment.

Insert Table 3 about here

Insert Figure 2 about here

3.3 ROI analysis

The ROI analysis revealed greater activity for the target sentence of *Ironic* than *Literal* stories in all four regions – the rTPJ, the lTPJ, the MPFC and the PC (see Figure 3). Specifically, paired t-tests revealed a significant increase of activity in the *Ironic > Literal* contrast for the rTPJ (t(19) = 2.42, *p* = .026), the MPFC (t(19) = 2.92, *p* = .009) and the PC (t(19) = 2.31, *p* = .032), whereas the lTPJ showed an activation increase that was marginally significant (t(19) = 1.98, *p* = .062).

Insert Figure 3 about here

To ensure that our results were not due to longer reading times in the *Ironic* than *Literal* condition, we also ran a model adding the reading times for the target sentence as covariates of no interest. The results obtained with this model (controlling for differences in reading times) did not alter the results obtained with our initial model. Specifically, all the clusters that were significantly more active in the *Ironic* versus *Literal* condition in the main analysis remained so when reading times were included as a covariate. That is, there was still

enhanced activity in the rTPJ (x = 62 y = -55 z = 29; Z = 4.42), lTPJ (x = -59 y = -43 z = 30; Z = 3.32), PC (x = 9 y = -60 z = 34; Z = 3.33) and MPFC (x = -6 y = 39 z = 40; Z = 5.23). This indicates that none of our results were due to differences in reading times when comparing *Ironic* and *Literal* conditions.

3.4 Psychophysiological interaction analysis

With respect to the *Ironic* condition compared to the *Literal* one, the PPI analysis revealed an increase in functional connectivity between the ventral part of the MPFC seed (x = 0 y = 55 z = 6) and the left IFG (peak: x = -50 y = 31 z = 0). There was also an increase of functional connectivity between the MPFC seed and the right IFG (peak: x = 53 y = 18 z = 25) (see Figure 4). Again, note that this occurs after we ensured that activity in the seed region was not dependent upon the type of stories (*Ironic* vs. *Literal*, see **Methods**). On the other hand, the PPI analysis using the Precuneus as seed region did not reveal any activation in any of the regions that were activated in the whole brain analysis for the contrast *Ironic > Literal*.

Insert Figure 4 about here

----

4. Discussion

This investigation was designed to determine the extent to which the Theory of Mind network is involved in the on-line processing of a pragmatically rich linguistic stimulus, if at all. We focused our investigation on ironic utterances because it is uncontroversial that such an utterance gives a listener access to a speaker's state of mind. Below, we review the results that demonstrate that Ironic utterances – when compared to their Literal controls -- do indeed activate the neural ToM network as well as prompt interactions with language areas. We also consider how these data impact ongoing debates and correspond with prior results.

Both the whole brain and ROI analyses conducted here show that understanding verbal irony engages a network of brain regions typically associated with Theory of Mind (the rTPJ, the lTPJ, the MPFC, and the PC; see for example Mitchell, 2009; Saxe, Carey, & Kanwisher, 2004; Van Overwalle & Baetens, 2009; also see Frith and Frith, 2006). *Ironic* target sentences – when compared to their *Literal* controls -- consistently elicited significant differences in each of four Regions of Interest that were chosen in a top-down manner and on the heels of an extensive and recent meta-analysis of the literature on ToM (Van Overwalle, 2009).

Importantly, our fMRI findings are unique when one takes into consideration the current neuroimaging literature on irony, which had shown very little ToM activity or none at all. We attribute the differences between our study and the prior seven to the materials and the presentation in that we aimed to bring the current study on irony in line with those found in the behavioral literature.

Critically, our investigation was also designed to determine whether or not irony processing would uncover evidence of integration between language and ToM processing and the results were positive. The PPI analysis showed that the functional connectivity between the ventral part of the MPFC and the left IFG increases when reading the target sentence in the *Ironic* condition when compared to the *Literal* one. Given that the vMPFC is crucial for ToM processing (e.g., Ma, Vandekerckhove, Van Overwalle, Seurinck, & Fias, 2011) and that the left IFG is strongly involved in the integration process in language (e.g., Hagoort, 2005), it is highly plausible there would be an exchange of information and integration precisely between these two regions. Importantly, the PPI analysis was carried out with seed regions that were chosen in a very conservative manner (i.e., the seed regions were not associated with the significant results reported above for the *Ironic* versus *Literal* contrasts). This arguably represents the strongest evidence yet that ToM is directly integrated with the language network *while* processing an utterance. Given that the procedure we used to determine inclusion into the PPI analysis was quite strict, the rTPJ and lTPJ were eliminated as potential seed regions. We are thus not in a position to determine their connectivity to language areas. More research needs to be done on this topic. For the moment, it is reasonable to suppose that the integration between different networks relies, not only on specific clusters of activation but also, on their patterns of connectivity.

As we said earlier, a growing body of literature associates coactivation of MPFC, rTPJ, lTPJ and PC with Theory of Mind processing (for reviews, see Mitchell, 2009; Saxe, Carey, & Kanwisher, 2004; Van Overwalle & Baetens, 2009; also see Frith and Frith, 2006). Therefore, the concomitant activation of these four regions in the present study is consistent with the hypothesis that irony processing involves ToM. Each of these regions, however, might cover a different aspect of what is more generally called "mentalizing activity". For example, it has been proposed that the rTPJ specifically supports the uniquely human ability to reason about the contents of mental states (Saxe, 2006). The MPFC, on the other hand, might be divided into two functional areas: the ventral part that might be implicated in emotional empathy, and

the more dorsal part that might be involved in representing the triadic relations between two minds and an object (a critical ability supporting shared attention and collaborative goals) (Saxe, 2006). Because the present study does not allow us to break down mentalizing activity into component parts, the question as to what role each ToM region plays in irony processing remains open. At present, we conclude only that irony processing simultaneously engages four regions typically associated with ToM.

Now that we have circumscribed the import of our data with respect to ToM, we would be remiss if we did not address the general psycholinguistic implications of our findings. We thus describe how these data speak to prior work on language processing in general and on irony-processing in particular by (i) considering how the ROIs can be viewed independently of ToM concerns, (ii) addressing potential criticisms from a psycholinguistic point of view and; (iii) considering how factors other than ToM and linguistic processing can have an influence on our results.

Although our findings are suggestive, we do not want to claim that the four regions on which we focus are *exclusively* dedicated to ToM processing. In fact, some specific aspects of language processing have been associated with these regions. For instance, the PC activations reported here can be viewed independently as a cluster that has been associated with integrating a sentence into context, in line with claims that link activity in the PC with situation model updating (e.g., Speer, Zacks, & Reynolds, 2007). Similarly, MPFC activations are often cited as evidence of general inference making (e.g., by Ferstl & von Cramon, 2002). In fact, it is not surprising that the results reported here are consistent with findings reported in the neurobiology of language. One can see the general consistency between our own findings with prior summaries by considering Ferstl and colleagues' model for language processing, the Extended Language Network (ELN), which highlights the regions that most frequently appear in studies on text comprehension (Ferstl, Neumann, Bogler, & Von Cramon, 2008). Besides the classic Broca's and Wernicke's areas, the network includes their right homologues, the middle and superior temporal lobes and, crucially for our purposes, the ToM regions (especially the dorso-medial prefrontal cortex).

One potential criticism of our study is that it is unbalanced because positive contexts are always followed by a positive target sentence whereas negative contexts are followed by either an ironic sentence or by one having a decoy (a positive and negative sentence,

respectively). The critique is that this leaves the target sentence in the *Literal* condition more predictable. Our response to this is threefold. First, the target sentences in the *Literal* condition were *designed* to be unremarkable and the source of floor effects so that they could be effective controls. If one comes up with unusual target sentences (e.g., a rare type of negative irony, which would provide a negative comment in light of a positive event) for the sake of having multiple controls, one could compromise the entire experiment. Consider Gibbs (1986) who – with similar reasoning -- included a control for an irony study in which a negative event was followed by a brutally honest remark (e.g., "You're not helping me" when a brother failed to show up to help on a construction project); the net result was that this control (a) provided the slowest reading times overall, (b) made *Literal* target sentences appear exceptional for their banality, which, in turn, (c) made *Literal* target latencies comparable to the *Ironic* ones. Our overall goal was to have ecologically valid remarks throughout that allowed for careful comparisons between the ironies and their controls. Second, there were no significant *Literal > Ironic* contrasts or any other indication that target sentences in the *Literal* condition were marked in some way. This indicates that there is nothing about the target sentences in the *Literal* condition that stands out when compared to those in the *Ironic* condition and that the Control worked as intended. Third, while the *Literal* condition is a convenient name for our study, we underline that from a participant's point of view, target sentences in this condition are simple, unmarked events similar to banal continuations. In the end, nearly half of banal continuations were read in the context of a negative event (as decoys) and the rest in the context of a positive one (*Literal* condition).

While the current study includes features that advance the neuroimaging literature on irony, while specifically underlining a role for ToM in irony processing, it rules out a potential cognitive confound too; namely, we showed how the study's irony-related effects persist even when reading times were included as a covariate. That said, there remain other extralinguistic issues that this study is not in a position to address. For instance, as indicated above, it would be relevant to know whether one would continue finding the results reported here when the ironic utterance itself is negative and its context is positive (e.g., "That was a failure" said upon completion of a successful event). On a similar front, it would be useful to rule out that ironic and the literal utterances yield differences because the implicit emotional valences for the ironic utterances (which are arguably negative) differ from their controls (which are arguably positive). Although we consider it likely that the reported ToM effects here would be maintained regardless of valences (e.g., see Morelli et al., 2012), this would need to be

established experimentally. In any case, we obviously do not want to claim that this single study is in a position to resolve all psycholinguistic and neural issues related to irony. Our modest but important goal from the start was to determine whether or not one can establish that ToM is intrinsically involved in irony processing; while this is a feature of irony that theoretical approaches take for granted it is a factor that experimental investigations have not detected until now.

Before concluding we address one last question: Do these results contradict the RH hypothesis? Given that we found activity in the RH, as expected, our response is necessarily negative. Nevertheless, our results do provide insights about the brain regions that are responsible for the deficits in irony processing in patients with RH lesions. If the rTPJ or the right part of MPFC is affected by lesions, then patients should be significantly impaired in their ability to infer others' mental states. This hypothesis calls for further investigation. Both lesion and TMS studies should help clarify the specific role that RH regions play in irony processing.

To summarize, we investigated verbal irony in order to determine the extent to which one finds activity in the ToM network as a result of a linguistic stimulus. Following a large number of studies on ToM processing, we focused on the rTPJ, the lTPJ, the MPFC and the precuneus and found greater activation in these regions for the *Ironic* condition when compared to the *Literal* condition. Moreover, analyses of functional connectivity suggest that information shared by the MPFC and with the left IFG is crucial for linking ToM activity with language processing. These results are in stark contrast with previous studies on verbal irony, which have found very little pointing to ToM activation. Although verbal irony represents perhaps one of the most obvious cases of pragmatic processing, we do not believe that evidence of ToM activity will be limited to such cases. It is our view that these findings on irony, based on ecologically valid materials, represent the tip of an iceberg. After all, one of the main tenets of pragmatics is that the linguistic code in all utterances underdetermines a speaker's meaning. It follows that all utterances require some amount of pragmatic processing in order to be understood by a listener. This points to a role for ToM. This study outlines how to go about finding ToM activity in language processing, which is by identifying circumstances that ought to engage a ToM network and by uncovering how the language network coordinates with it during the on-line processing of an utterance.

5. Acknowledgements

6. References

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*(3-4), 203–216.

Brownell, H. H., Carroll, J. J., Rehak, A., & Wingfield, A. (1992). The use of pronoun anaphora and speaker mood in the interpretation of conversational utterances by right hemisphere brain-damaged patients. *Brain and language*, *43*(1), 121–47.

Carston, R. (2002). *Thoughts and Utterances: The Pragmatics of Explicit Communication*. Oxford: Blackwell.

Chevallier, C., Kohls, G., Troiani, V., Brodkin, E. S., & Schultz, R. T. (2012). The social motivation theory of autism. *Trends in cognitive science*, *16*(4), 231–239.

Chevallier, C., Noveck, I. A., Happé, F., & Deirdre, W. (2011). What's in a voice? Prosody as a test case for the Theory of Mind account of autism. *Neuropsychologia*, *49*(3), 507–517.

Chevallier, C., & Wilson, D. (2010). Scalar inferences in autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *40*(9), 1104–1117.

Clark, H. H., & Gerrig, R. J. (1984). On the Pretense Theory of Irony, *Journal of Experimental Psychology: General*, *113*(1), 121–126.

Eviatar, Z., & Just, M. A. (2006). Brain correlates of discourse processing: An fMRI investigation of irony and conventional metaphor comprehension. *Neuropsychologia*, *44*(12), 2348–2359.

Ferstl, E. C., Neumann, J., Bogler, C., & Von Cramon, D. Y. (2008). The extended language network: A meta-analysis of neuroimaging studies on text comprehension. *Human brain mapping*, *29*(5), 581–593.

Friston, K. J., Buechel, C., Fink, G., Morris, J., Rolls, E., & Dolan, R. (1997). Psychophysiological and Modulatory Interactions in Neuroimaging. *Neuroimage*, *6*(3), 218–229.

Frith, C. D., & Frith, U. (2006). The Neural Basis of Mentalizing Minireview. *Neuron*, *50*(4), 531–534.

Gernsbacher, M. A., & Robertson, D. A. (2004). Watching the brain comprehend discourse. In A. Healy (Ed.), *Experimental cognitive psychology and its applications* (pp. 157–67). Washington, DC: APA Publications.

Gibbs, R. W. (1986). On the psycholinguistics of sarcasm. *Journal of Experimental Psychology: General*, *115*(1), 3–15.

Gibbs, R. W. (2000). Irony in talk among friends. *Metaphor and Symbol*, *15*(1-2), 5–27.

Gitelman, D. R., Penny, W. D., Ashburner, J., & Friston, K. J. (2003). Modeling regional and psychophysiologic interactions in fMRI: the importance of hemodynamic deconvolution. *Neuroimage*, *19*(1), 200–207.

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

Hagoort, P. (2005). On Broca, brain, and binding: a new framework. *Trends in cognitive sciences*, *9*(9), 416–23.

Happé, F. G. (1993). Communicative competence and theory of mind in autism: a test of relevance theory. *Cognition*, *48*(2), 101–19.

Josephs, O., Turner, R., & Friston, K. J. (1997). Event-related fMRI. *Human brain mapping*, *5*(4), 243–248.

Kreuz, R. J., Roberts, R. M., Johnson, B. K., & Bertus, L. E. (1996). Figurative language occurrence and co-occurence in contemporary literature. In R. J. Kreuz & M. S. MacNealy (Eds.), *Empirical approaches to literature and aesthetics* (pp. 83–97). Norwood, NJ: Ablex.

Long, D. L., & Baynes, K. (2002). Discourse representation in the two cerebral hemispheres. *Journal of cognitive neuroscience*, *14*(2), 228–42.

Ma, N., Vandekerckhove, M., Van Overwalle, F., Seurinck, R., & Fias, W. (2011). Spontaneous and intentional trait inferences recruit a common mentalizing network to a

different degree: spontaneous inferences activate only its core areas. *Social neuroscience*, *6*(2), 123–38.

Malle, B. (2002). The relation between language and theory of mind in development and evolution. In T. Givon & B. F. Malle (Eds.), *The evolution of language out of pre-language* (pp. 265–284). Am: John Benjamins B. V.

McDonald, S. (1999). Exploring the process of inference generation in sarcasm: a review of normal and clinical studies. *Brain and language*, *68*(3), 486–506.

Mitchell, J. P. (2009). Inferences about mental states. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, *364*(1521), 1309–16.

Mitchell, J. P., Macrae, C. N., & Banaji, M. R. (2006). Dissociable medial prefrontal contributions to judgments of similar and dissimilar others. *Neuron*, *50*(4), 655–63.

Morelli, S.A., Rameson, L., & Lieberman, M.D., April 2012. Empathy for positive and negative emotions: Implications for brain and behavior. Talk to be presented as symposium chair at the 5th Annual Social Affective Neuroscience Conference. New York, NY

Overwalle, F. V., & Baetens, K. (2009). Understanding others' actions and goals by mirror and mentalizing systems: a meta-analysis. *Neuroimage*, *48*(3), 564–584. d

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind. *Behavioral and brain sciences*, *1*(4), 515–526.

Rapp, A. M., Leube, D. T., Erb, M., Grodd, W., & Kircher, T. T. J. (2007). Laterality in metaphor processing: Lack of evidence from functional magnetic resonance imaging for the right hemisphere theory. *Brain and Language*, *100*(2), 142–149.

Rapp, A. M., Mutschler, D., Wild, B., Erb, M., Lengsfeld, I., Saur, R., & Grodd, W. (2010). Neural correlates of irony comprehension: The role of schizotypal personality traits. *Brain and language*, *113*(1), 1–12. doi:10.1016/j.bandl.2009.11.007

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*(2), 277–293.

Saxe, R. (2006). Uniquely human social cognition. *Current opinion in neurobiology*, *16*(2), 235–9.

Saxe, R., Carey, S., & Kanwisher, N. (2004). Understanding other minds: linking developmental psychology and functional neuroimaging. *Annual review of psychology*, *55*, 87–124.

Saxe, R., & Powell, L. J. (2006). It's the thought that counts: specific brain regions for one component of theory of mind. *Psychological science*, *17*(8), 692-699.

Shamay-Tsoory, S. G., Tomer, R., & Aharon-Peretz, J. (2005). The neuroanatomical basis of understanding sarcasm and its relationship to social cognition. *Neuropsychology*, *19*(3), 288–300.

Shibata, M., Toyomura, A., Itoh, H., & Abe, J. (2010). Neural substrates of irony comprehension: A functional MRI study. *Brain research*, *1308*, 114–123.

Speer, N. K., Zacks, J. M., & Reynolds, J. R. (2007). Human brain activity time-locked to narrative event boundaries. *Psychological science*, *18*(5), 449–55.

Sperber, D., & Wilson, D. (1986). *Relevance: Communication and Cognition*. Cambridge, Mass.: Oxford : Blackwell; Wiley-Blackwell; 2nd Edition edition 1995.

Tager-Flusberg, H. (2000). Language and Understanding Minds in Autism. In S. Baron-Cohen, H. Tager-Flusberg, & D. J. Cohen (Eds.), *Understanding other minds: Perspectives from autism and developmental cognitive neuroscience* (Second., pp. 124 – 149). Oxford: Oxford University Press.

Talairach, J., & Tournoux, P. (1988). *Co-planar stereotaxic atlas of the human brain: 3-dimensional proportional system: an approach to cerebral imaging*. New York: Thieme.

Uchiyama, H. T., Saito, D. N., Tanabe, H. C., & Harada, T. (2011). Distinction between the literal and intended meanings of sentences: A functional magnetic resonance imaging study of metaphor and sarcasm. *Cortex*, *48*(5), 563–583.

Uchiyama, H. T., Seki, A., Kageyama, H., Saito, D. N., Koeda, T., Ohno, K., & Sadato, N. (2006). Neural substrates of sarcasm: a functional magnetic-resonance imaging study. *Brain research*, *1124*(1), 100–10.

Wakusawa, K., Sugiura, M., Sassa, Y., Jeong, H., Horie, K., Sato, S., Yokoyama, H., et al. (2007). Comprehension of implicit meanings in social situations involving irony: A functional MRI study. *NeuroImage*, *37*(4), 1417–1426.

Wang, A. T., Lee, S. S., Sigman, M., & Dapretto, M. (2006). Developmental changes in the neural basis of interpreting communicative intent. *Social cognitive and affective neuroscience*, *1*(2), 107–121.

Wilson, D. (2006). The pragmatics of verbal irony: Echo or pretence? *Lingua*, *116*(10), 1722–1743.

Wilson, D., & Sperber, D. (1992). On verbal irony. *Lingua*, *87*(1), 53–76.

Wilson, Deirdre. (2009). Irony and metarepresentation. *UCLWPL*, *21*, 183–226.

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, *13*(1), 103–128.

Appendix A: Experimental stimuli

Condition	French (as presented)	English translations
Ironic	Cynthia et Léa chantent dans le même opéra. Le soir de la première, elles se retrouvent au théâtre. Le spectacle commence pile à l'heure. Durant la représentation, elles font beaucoup de fausses notes. Après le spectacle, Cynthia dit à Léa : « Ce soir on a fait une performance magistrale. » Tandis qu'elles se démaquillent, les deux filles continuent à parler du spectacle. Question : À votre avis, est-ce que la performance est le matin ?	Cynthia and Léa sing together in the same opera. On the night of the premiere they meet at the theatre. The show begins exactly on time. During their performance they often sing off key. After the show, Cynthia says to Léa: "Tonight we gave a superb performance." As they take off their make-up they continue to discuss the show. Question: In your opinion, do you think that the performance was in the morning?
Literal	Cynthia et Léa chantent dans le même opéra. Le soir de la première, elles se retrouvent au théâtre. Le spectacle commence pile à l'heure. La représentation est excellente et les chanteurs sont longuement applaudis. Après le spectacle, Cynthia dit à Léa : « Ce soir on a fait une performance magistrale. » Tandis qu'elles se démaquillent, les deux filles continuent à parler du spectacle. Question : À votre avis, est-ce que la performance est le matin ?	Cynthia and Léa sing together in the same opera. On the night of the premiere they meet at the theatre. The show begins exactly on time. The show was excellent and the singers were given a long applause. After the show, Cynthia says to Léa: "Tonight we gave a superb performance." As they take off their make-up they continue to discuss the show. Question: In your opinion, do you think that the performance was in the morning?
Ironic	Lors d'un dîner, Patrick parle à son collègue Pascal de sa fondation contre l'anorexie. Il explique à Pascal sa nouvelle idée. Il lui dit qu'il va lancer une nouvelle campagne de financement. Quelques mois après ils se retrouvent pour évaluer les résultats décevants de la campagne. Les gens ont donné beaucoup moins cette fois-ci et Pascal dit à Patrick : « Avec cette campagne on a fait un grand coup. » Patrick et Pascal commencent à penser aux nouvelles activités pour la fondation. Question : A votre avis, est-ce que Pascal et Patrick sont collègues ?	While at dinner, Patrick talks to his colleague Pascal about his foundation to combat anorexia. He explains to Patrick his new idea. He tells him that he is going to begin a new fundraising campaign. Several months later, they meet again to evaluate the disappointing results of the campaign. People gave much less this time and Pascal says to Patrick: "This campaign has really been a hit." Patrick and Pascal start to think of new activities for the foundation. Question: In your opinion, do you think Pascal and Patrick are colleagues?
Literal	Lors d'un dîner, Patrick parle à son collègue Pascal de sa fondation contre l'anorexie. Il explique à Pascal sa nouvelle idée. Il lui dit qu'il va lancer une nouvelle campagne de financement. Quelques mois après ils se retrouvent pour évaluer les excellents résultats de la campagne. Les gens ont donné beaucoup plus que les années précédentes et Pascal dit à Patrick : « Avec cette campagne on a fait un grand coup ! » Patrick et Pascal commencent à penser aux nouvelles	While at dinner, Patrick talks to his colleague Pascal about his foundation to combat anorexia. He explains to Patrick his new idea. He tells him that he is going to begin a new fundraising campaign. Several months later, they meet again to evaluate the excellent results of the campaign. People gave much more than in previous years and Pascal says to Patrick: "This campaign has really been a hit!" Patrick and Pascal start to think of new activities for the

	activités pour la fondation. Question : A votre avis, est-ce que Pascal et Patrick sont collègues ?	foundation. Question: In your opinion, do you think Pascal and Patrick are colleagues?
Ironic	Luc parle de ses investissements à Alfred qui est courtier en bourse. Luc voudrait investir son argent dans les actions d'une petite entreprise. Alfred lui explique les avantages et les inconvénients. Luc achète les actions mais un mois plus tard leur valeur a diminué de moitié. Lors de la réunion suivante, Luc en reparle à Alfred et dit : <u>« C'est ce qui s'appelle un investissement rentable. »</u> Pendant qu'ils parlent, les nouveaux cours de la bourse s'affichent sur le portable de Luc. Question : A votre avis, est-ce que Luc veut investir dans une grande entreprise ?	Luc talks about an investment with Alfred, who is a stockbroker. Luc is interested in buying shares of a small company. Alfred describes the advantages and inconveniences of such an investment. Luc buys the stocks, but one month later their value has dropped by half. At the next meeting, Luc talks about it with Alfred again and says: <u>"This is what's called a worthwhile investment."</u> As they are talking, Luc's laptop displays new stock quotes. Question: In your opinion, does Luc want to invest in a big company?
Literal	Luc parle de ses investissements à Alfred qui est courtier en bourse. Luc voudrait investir son argent dans les actions d'une petite entreprise. Alfred lui explique les avantages et les inconvénients. Un mois plus tard, le cours des actions de Luc a déjà doublé. Au cours d'un déjeuner, il en reparle à Alfred : <u>« C'est ce qui s'appelle un investissement rentable. »</u> Pendant qu'ils parlent, les nouveaux cours de la bourse s'affichent sur le portable de Luc. Question : A votre avis, est-ce que Luc veut investir dans une grande entreprise ?	Luc talks about an investment with Alfred, who is a stockbroker. Luc is interested in buying shares of a small company. Alfred describes the advantages and inconveniences of such an investment. One month later their value, Luc's stocks have already doubled. Over a lunch, Luc talks about it again with Alfred: <u>"This is what's called a worthwhile investment."</u> As they are talking, Luc's laptop displays new stock quotes. Question: In your opinion, does Luc want to invest in a big company?
Ironic	Clara et Isabelle doivent décider quel film aller voir au cinéma. Elles remarquent l'affiche d'un film dans la rue. Elles ne le connaissent pas mais décident d'aller le voir. Les deux amies achètent les billets et des pop-corn. Le film se révèle être banal et très ennuyeux, Clara dit alors à Isabelle : <u>« Nous sommes allées voir un film formidable. »</u> Elles sortent de la salle et vont s'acheter une glace. Question : A votre avis, est-ce que Clara et Isabelle vont s'acheter une glace ?	Clara and Isabelle must decide which film to see at the cinema. They see a poster for a film outside. They aren't familiar with it but they decide to go see it. The two friends buy tickets and popcorn. The film turns out to be banal and very boring, so Clara says to Isabelle: <u>"We went to see a wonderful film."</u> They leave the theater and go buy an ice cream. Question: In your opinion, do you think Clara and Isabelle went to buy an ice cream?
Literal	Clara et Isabelle doivent décider quel film aller voir au cinéma. Elles remarquent l'affiche d'un film dans la rue. Elles ne le connaissent pas mais décident d'aller le voir. Les deux amies achètent les billets et des pop-corn. Le film se révèle être excitant et surprenant, Clara dit alors à Isabelle : <u>« Nous sommes allées voir un film formidable. »</u> Elles sortent de la salle et vont s'acheter une glace. Question : A votre avis, est-ce que Clara et Isabelle vont s'acheter une glace ?	Clara and Isabelle must decide which film to see at the cinema. They see a poster for a film outside. They aren't familiar with it but they decide to go see it. The two friends buy tickets and popcorn. The film turns out to be exciting and surprising, so Clara says to Isabelle: <u>"We went to see a wonderful film."</u> They leave the theater and go buy an ice cream. Question: In your opinion, do you think Clara and Isabelle went to buy an ice cream?
Ironic	Léonard propose à son nouveau collègue Gustave de lui montrer la ville. Ils veulent commencer par la place principale. La ville est grande et ils décident de se rejoindre tôt dans l'après-midi. Il se met à pleuvoir et ils sont forcés de rester au café sans rien voir de la ville. Le soir Léonard dit à Gustave : <u>« Nous avons fait une visite fantastique. »</u> Les deux collègues se donnent alors rendez-vous le lendemain au bureau.	Léonard offers to show his new co-worker Gustave around town. They want to start at the main square. The city is big and they decide to meet up early in the afternoon. It begins to rain and they are forced to stay at a café without seeing any of the city. That evening, Léonard says to Gustave: <u>"We've carried out a fantastic tour."</u> The two colleagues then plan to meet again the next day at work.

	Question : A votre avis, est-ce que Gustave habite dans cette ville depuis de nombreuses années ?	Question: In your opinion, do you think Gustave has lived in that city for many years?
Literal	Léonard propose à son nouveau collègue Gustave de lui montrer la ville. Ils décident de commencer par la place principale. La ville est grande et ils décident de se rejoindre tôt dans l'après-midi. Ils marchent tout l'après-midi et découvrent toutes les merveilles de la ville. Le soir Léonard dit à Gustave : « Nous avons fait une visite fantastique. » Les deux collègues se donnent alors rendez-vous le lendemain au bureau. Question : A votre avis, est-ce que Gustave habite dans cette ville depuis de nombreuses années ?	Léonard offers to show his new co-worker Gustave around town. They decide to start at the main square. The city is big and they decide to meet up early in the afternoon. They walk around all afternoon and see all the sights of the city. That evening, Léonard says to Gustave: "We've carried out a fantastic tour." The two colleagues then plan to meet again the next day at work. Question: In your opinion, do you think Gustave has lived in that city for many years?
Ironic	Olivier a pris du poids et décide d'aller chez son médecin. Ce dernier lui prescrit un régime. Olivier regarde les recettes et achète tout le nécessaire pour suivre le régime. Un mois après le début du régime, il n'a même pas perdu un gramme. Olivier dit alors à son médecin : « Ce régime me donne beaucoup d'espoir. » Olivier prend un nouveau rendez-vous pour contrôler son poids un mois plus tard. Question : A votre avis, Olivier refuse-t-il de suivre le régime ?	Olivier has gained weight and he decides to go see his doctor. The doctor puts him on a diet. Olivier looks at the recipes and buys everything you need to follow the diet. One month after starting the diet, he has not even lost one gram. So Olivier tells his doctor: "This diet really gives me hope." Olivier makes a new appointment to check on his weight a month later. Question: In your opinion, does Oliver refuse to follow the diet?
Literal	Olivier a pris du poids et décide d'aller chez son médecin. Ce dernier lui prescrit un régime. Olivier regarde les recettes et achète tout le nécessaire pour suivre le régime. Un mois après le début du régime, il a déjà perdu sept kilogrammes. Olivier dit alors à son médecin : « Ce régime me donne beaucoup d'espoir. » Olivier prend un nouveau rendez-vous pour contrôler son poids un mois plus tard. Question : A votre avis, Olivier refuse-t-il de suivre le régime ?	Olivier has gained weight and he decides to go see his doctor. The doctor puts him on a diet. Olivier looks at the recipes and buys everything you need to follow the diet. One month after starting the diet, he has already lost seven kilograms. So Olivier tells his doctor: "This diet really gives me hope." Olivier makes a new appointment to check on his weight a month later. Question: In your opinion, does Oliver refuse to follow the diet?
Ironic	En sortant d'une conférence à l'étranger, Benoît croise son collègue Thierry. Il lui demande où se trouve la cafétéria de l'université. Thierry lui propose de l'accompagner et Benoît accepte. Après de longues recherches ils arrivent enfin à la cafétéria. Benoît dit : « On a trouvé la cafétéria rapidement. » Les places sont toutes prises, il y a déjà beaucoup de monde à l'intérieur. Question : A votre avis, est-ce que la cafétéria est ouverte ?	While leaving a conference abroad, Thierry runs into his colleague Benoît. He asks him where the university cafeteria is. Thierry offers to accompany him and Benoît accepts. After a long search they finally arrive at the cafeteria. Benoît says: "We found the cafeteria quickly." There are already a lot of people inside, and all the seats are taken. Question: In your opinion, is the cafeteria open?
Literal	En sortant d'une conférence à l'étranger, Benoît croise son collègue Thierry. Il lui demande où se trouve la cafétéria de l'université. Thierry lui propose de l'accompagner et Benoît accepte. Ils trouvent la cafétéria tout de suite, elle est au coin de la rue. Benoît dit : « On a trouvé la cafétéria rapidement. » Les places sont toutes prises, il y a déjà beaucoup de monde à l'intérieur. Question : A votre avis, est-ce que la cafétéria est ouverte ?	While leaving a conference abroad, Thierry runs into his colleague Benoît. He asks him where the university cafeteria is. Thierry offers to accompany him and Benoît accepts. They find the cafeteria right away; it is just around the corner. Benoît says: "We found the cafeteria quickly." There are already a lot of people inside, and all the seats are taken. Question: In your opinion, is the cafeteria open?
Ironic	Claude rencontre son collègue Steve sur la route qui mène à l'université. Ils ont le même cours ce matin. Ils s'installent et écoutent le professeur. Les deux étudiants trouvent le cours difficile, sans intérêt et	Claude meets his colleague Steve on the road leading up to their university. They have the same class this morning. They settle in and listen to the professor. The two students find the lecture difficult, uninteresting and

	soporifique. A la fin Claude dit à Steve : « Aujourd'hui, le professeur était captivant. » Fatigués, les deux étudiants vont boire un café avant le cours suivant. Question : A votre avis, Claude et Steve suivent-ils des cours différents ?	tiresome. At the end Claude says to Steve: "The professor was in captivating today." Tired, the two students go to drink a coffee before their next class. Question: In your opinion, do Claude and Steve take different courses?
Literal	Claude rencontre son collègue Steve sur la route qui mène à l'université. Ils ont le même cours ce matin. Ils s'installent et écoutent le professeur. Le cours se révèle être très intéressant et stimulant pour les deux étudiants. A la fin Claude dit à Steve : « Aujourd'hui, le professeur était captivant. » Fatigués, les deux étudiants vont boire un café avant le cours suivant. Question : A votre avis, Claude et Steve suivent-ils des cours différents ?	Claude meets his colleague Steve on the road leading up to their university. They have the same class this morning. They settle in and listen to the professor. The presentation proved to be very interesting and stimulating for both students. At the end Claude says to Steve: "The professor was captivating today." Tired, the two students go to drink a coffee before their next class. Question: In your opinion, do Claude and Steve take different courses?
Ironic	Josiane est malade et demande à son colocataire Félix de lui préparer un remède aux plantes. Josiane doit rester au lit et se reposer toute la journée. Son colocataire lui prépare sa boisson aux plantes. Après l'avoir bu, Josiane se sent encore plus mal et finit par vomir. Son colocataire voit que son état a empiré et lui dit : « Ce remède a montré son efficacité. » Le soir, Josiane et son colocataire regardent un film à la télé. Question : A votre avis, est-ce que Josiane est malade ?	Josiane is ill and asks her roommate to prepare her an herbal remedy. Josiane must stay in bed and rest all day. Her roommate makes her the herbal drink. After drinking it, Josiane feels even worse and ends up vomiting. Her roommate sees that she is even sicker and says to her: "This remedy has really shown its effectiveness." That evening, Josiane and her roommate watch a movie on television. Question: In your opinion, is Josiane sick?
Literal	Josiane est malade et demande à son colocataire Félix de lui préparer un remède aux plantes. Josiane doit rester au lit et se reposer toute la journée. Son colocataire lui prépare sa boisson aux plantes. Peu de temps après l'avoir bu, Josiane se sent beaucoup mieux. Son colocataire voit que son état s'est amélioré et lui dit : « Ce remède a montré son efficacité. » Le soir, Josiane et son colocataire regardent un film à la télé. Question : A votre avis, est-ce que Josiane est malade ?	Josiane is ill and asks her roommate to prepare her an herbal remedy. Josiane must stay in bed and rest all day. Her roommate makes her the herbal drink. Shortly after drinking it, Josiane feels much better. Her roommate sees that she feels better and says to her: "This remedy has really shown its effectiveness." That evening, Josiane and her roommate watch a movie on television. Question: In your opinion, is Josiane sick?
Ironic	Tom et Dave parlent des élections du conseil d'administration de la faculté. Dave est représentant des étudiants et sait que les résultats sont importants. Le vote a commencé il y a quelques minutes. Après plusieurs heures, ils apprennent que le parti du président actuel a perdu. Tom dit à Dave : « Le président doit sans doute être ravi. » Tom et Dave continuent de parler des élections durant un bon moment. Question : A votre avis, est-ce que Tom et Dave sont membres du conseil d'administration ?	Tom and Dave are talking about the elections for the university's board of directors. Dave is the student representative and knows that the results are important. The voting began a few minutes ago. After several hours they learn that the current president's party lost. Tom tells Dave: "The president must certainly be delighted." Tom and Dave continue to talk about the elections for a while. Question: In your opinion, are Tom and Dave members of the board?
Literal	Tom et Dave parlent des élections dans le conseil d'administration de la faculté. Dave est représentant des étudiants et sait que les résultats sont importants. Le vote a commencé il y a quelques minutes. Après plusieurs heures, ils apprennent que le parti du président actuel a largement gagné. Tom dit à Dave : « Le président doit sans doute être ravi. » Tom et Dave continuent de parler des élections durant un bon moment. Question : A votre avis, est-ce que Tom et Dave sont	Tom and Dave are talking about the elections for the university's board of directors. Dave is the student representative and knows that the results are important. The voting began a few minutes ago. After several hours they learn that the current president's party won decisively. Tom tells Dave: "The president must certainly be delighted." Tom and Dave then continue to talk about the elections for a while.

	membres du conseil d'administration ?	Question: In your opinion, are Tom and Dave members of the board?
Ironic	Michel et Jim pêchent ensemble une fois par an dans un lac. Comme d'habitude, ils discutent en attendant que les poissons mordent à l'hameçon. Ils se racontent leurs aventures de pêche. A la fin de la journée, aucun d'entre eux n'a réussi à attraper un seul poisson. En partant, Michel dit à Jim : « Cette journée a été très productive. » Michel et Jim rentrent chez eux profitant des derniers rayons du soleil. Question : A votre avis, est-ce que Michel et Jim rentrent chez eux dans la soirée ?	Michel and Jim go fishing together once a year in a lake. As usual, they talk while waiting for the fish to bite. They talk about their fishing adventures. At the end of the day, neither of them managed to catch a single fish. As they were leaving, Michel told Jim: "This has been a really productive day." Michel and Jim return home with the last rays of the sun. Question: In your opinion, do Michel and Jim return home in the evening?
Literal	Michel et Jim pêchent ensemble une fois par an dans un lac. Comme d'habitude, ils discutent en attendant que les poissons mordent à l'hameçon. Ils se racontent leurs aventures de pêche. Aujourd'hui, les deux pêcheurs ont rempli entièrement leurs nasses à poissons. En partant, Michel dit à Jim : « Cette journée a été très productive. » Michel et Jim rentrent chez eux profitant des derniers rayons du soleil. Question : A votre avis, est-ce que Michel et Jim rentrent chez eux dans la soirée ?	Michel and Jim go fishing together once a year in a lake. As usual, they talk while waiting for the fish to bite. They talk about their fishing adventures. Today, the two fishermen have completely filled their nets with fish. As they were leaving, Michel told Jim: "This has been a really productive day." Michel and Jim return home with the last rays of the sun. Question: In your opinion, do Michel and Jim return home in the evening?
Ironic	Cédric et Bob se retrouvent dans le même avion pour New York. Tous les deux prennent très souvent l'avion. Pendant le vol, ils parlent de leur voyage professionnel. A cet instant, on leur annonce que leur arrivée sera retardée. Cédric dit à Bob : « Je suis impressionné par leur ponctualité. » Une fois arrivés, ils partent tout de suite à la recherché d'un taxi libre. Question : A votre avis, est-ce la première fois que Bob prend l'avion ?	Cédric and Bob find themselves on the same plane to New York. Both of them fly often. During the flight, they talk about their business trip. Just then, the pilot announces that their arrival will be delayed. Cédric says to Bob: "I am impressed by their punctuality." Once there, they go immediately in search of a taxi. Question: In your opinion, is this the first time Bob rides an airplane?
Literal	Cédric et Bob se retrouvent dans le même avion pour New York. Tous les deux prennent très souvent l'avion. Pendant le vol, ils parlent de leur voyage professionnel. A cet instant, on leur annonce que leur avion sera à l'heure. Cédric dit à Bob : « Je suis impressionné par leur ponctualité. » Une fois arrivés, ils partent tout de suite à la recherché d'un taxi libre. Question : A votre avis, est-ce la première fois que Bob prend l'avion ?	Cédric and Bob find themselves on the same plane to New York. Both of them fly often. During the flight, they talk about their business trip. Just then, the pilot announces that their plane will land on time. Cédric says to Bob: "I am impressed by their punctuality." Once there, they go immediately in search of a taxi. Question: In your opinion, is this the first time Bob rides an airplane?
Ironic	Karen et Jean-Claude jouent aux échecs dans un parc. Ils viennent juste de commencer leur partie. Chacun élabore sa propre stratégie d'attaque. Après seulement six coups, Karen parvient à battre Jean-Claude. Jean-Claude lui dit alors : « C'est clair, nous avons tout à fait le même niveau. » Karen et Jean-Claude programment un autre rendez-vous pour rejouer ensemble. Question : A votre avis, est-ce que Karen et Jean-Claude aiment jouer aux échecs ?	Karen and Jean-Claude are playing chess in a park. They have just started their game. Each plans out their own strategy of attack. After only six moves, Karen manages to beat Jean-Claude. So Jean-Claude says to her: "Clearly, we have exactly the same level." Karen and Jean-Claude schedule a new appointment to play together again. Question: In your opinion, do Jean-Claude and Karen like to play chess?
Literal	Karen et Jean-Claude jouent aux échecs dans un parc. Ils viennent juste de commencer leur partie. Chacun élabore sa propre stratégie d'attaque. Après deux heures de jeu, personne n'a pris l'avantage.	Karen and Jean-Claude are playing chess in a park. They have just started their game. Each develops its own strategy of attack. After two hours of play, no one had taken the lead.

	Jean-Claude lui dit alors : « C'est clair, nous avons tout à fait le même niveau. » Karen et Jean-Claude programment un autre rendez-vous pour rejouer ensemble. Question : A votre avis, est-ce que Karen et Jean-Claude aiment jouer aux échecs ?	So Jean-Claude says to her: "Clearly, we have exactly the same level." Karen and Jean-Claude schedule a new appointment to play together again. Question: In your opinion, do Jean-Claude and Karen like to play chess?
Ironic	Joan et Marc sont deux critiques gastronomiques qui doivent évaluer un restaurant. Installés à leur table, ils attendent d'être servis. Entre temps, ils discutent du travail. Le serveur n'arrive qu'après un long moment et accumule les erreurs durant toute la soirée. Lorsque le serveur apporte l'addition à la fin du repas, Joan dit à Marc : « On ajoutera deux points pour le service. » Après l'inspection du restaurant, Joan et Marc reprennent leur examen. Question : A votre avis, est-ce que Joan et Marc vont donner leur avis sur le restaurant ?	Joan and Marc are two food critics who must evaluate a restaurant. Seated at their table, they are waiting to be served. Meanwhile, they discuss their job. The server takes a long time to arrive and accumulates errors throughout the evening. When the waiter brings the bill at the end of the meal, Joan says to Marc: "We'll add two points for the service." After inspecting the restaurant, Joan and Mark resume their examination. Question: In your opinion, are Joan and Marc going to give their opinion about the restaurant?
Literal	Joan et Marc sont deux critiques gastronomiques qui doivent évaluer un restaurant. Installés à leur table, ils attendent d'être servis. Entre temps, ils discutent du travail. Ils sont servis rapidement et le serveur reste très attentionné pendant toute la soirée. Lorsque le serveur apporte l'addition à la fin du repas, Joan dit à Marc : « On ajoutera deux points pour le service. » Après l'inspection du restaurant, Joan et Marc reprennent leur examen. Question : A votre avis, est-ce que Joan et Marc vont donner leur avis sur le restaurant ?	Joan and Mar are two food critics who must evaluate a restaurant. Seated at their table, they are waiting to be served. Meanwhile, they discuss their job. They are served quickly and the server remains very attentive throughout the evening. When the waiter brings the bill at the end of the meal, Joan says to Marc: "We'll add two points for the service." After inspecting the restaurant, Joan and Mark resume their examination. Question: In your opinion, are Joan and Marc going to give their opinion about the restaurant?
Ironic	Hélène et Gérard sont à une vente aux enchères de tableaux. Gérard, qui est amateur d'art, explique à Hélène la valeur des toiles présentées. Hélène est très intéressée et l'écoute attentivement. A ce moment, un tableau d'art moderne est proposé mais personne ne surenchérit. Hélène dit alors à Gérard : « La concurrence est rude pour ce tableau. » Avant de partir, Gérard parvient à conclure quelques très bonnes affaires. Question : A votre avis, est-ce que Hélène et Gérard sont à une exposition d'art ?	Helen and Gerard are at an art auction. Gerard, who is an art lover, explains to Helen the value of the paintings presented. Helen is very interested and listens to him carefully. A modern art painting is presented, but no one bids on it. So Helen says to Gerard: "The competition is stiff for this painting." Before leaving, Gerard manages to get some good deals. Question: In your opinion, are Helen and Gerard at an art exhibition?
Literal	Hélène et Gérard sont à une vente aux enchères de tableaux. Gérard, qui est amateur d'art, explique à Hélène la valeur des toiles présentées. Hélène est très intéressée et l'écoute attentivement. A ce moment, un tableau d'art moderne est très prisé et les enchères montent très haut. Hélène dit alors à Gérard : « La concurrence est rude pour ce tableau. » Avant de partir, Gérard parvient à conclure quelques très bonnes affaires. Question : A votre avis, est-ce que Hélène et Gérard sont à une exposition d'art ?	Helen and Gerard are at an art auction. Gerard, who is an art lover, explains to Helen the value of the paintings presented. Helen is very interested and listens to him carefully. One modern art painting was very popular and the bidding went very high. So Helen says to Gerard: "The competition is stiff for this painting." Before leaving, Gerard manages to get some good deals. Question: In your opinion, are Helen and Gerard at an art exhibition?
Ironic	Armelle et Sally partent en voyage à la campagne avec une vieille voiture. Des amis les attendent le soir même dans une maison pour dîner. Durant le voyage, elles discutent de leur soirée à venir. Après une heure de trajet, la voiture tombe en panne. Sally dit à Armelle : « Comme ça c'est sûr qu'on arrivera à l'heure. » A cet instant, leurs amis les appellent pour savoir où elles se	Armelle and Sally leave for a trip to the countryside in an old car. Their friends wait for them to arrive at the house for dinner that evening. During the trip, they discuss their evening to come. After an hour's drive, the car breaks down. Sally says to Armelle: "At this rate we're sure to get there on time." At that moment, their friends call them to find out where they

	trouvent.	are.
	Question : A votre avis, est-ce qu'Armelle a une vieille voiture ?	Question: In your opinion, does Armelle have an old car?
Literal	Armelle et Sally partent en voyage à la campagne avec une vieille voiture. Des amis les attendent le soir même dans une maison pour dîner. Durant le voyage, elles discutent de leur soirée à venir. Au milieu de l'après-midi, elles sont déjà presque arrivées. Sally dit à Armelle : « Comme ça c'est sûr qu'on arrivera à l'heure. » A cet instant, leurs amis les appellent pour savoir où elles se trouvent. Question : A votre avis, est-ce qu'Armelle a une vieille voiture ?	Armelle and Sally leave for a trip to the countryside in an old car. Their friends wait for them to arrive at the house for dinner that evening. During the trip, they discuss their evening to come. By mid-afternoon, they are already almost there. Sally says to Armelle: "At this rate we're sure to get there on time." At that moment, their friends call them to find out where they are. Question: In your opinion, does Armelle have an old car?
Ironic	Georges fait la lessive avec son colocataire Yves. Yves fait remarquer que Georges a beaucoup de chemises à laver. Ils passent un peu de temps à chercher le bon programme et la bonne température. Au moment d'étendre le linge, ils voient que les chemises sont encore pleines de tâches. Georges dit à Yves : « Cette machine fait un lavage impeccable. » Après avoir étendu le linge, ils vont boire une bière dans la cuisine. Question : A votre avis, est-ce que Georges et Yves vont dans un pub ?	Georges is doing the laundry with his roommate Yves. Yves points out that Georges has a lot of shirts to wash. They spend some time looking for the right program and the right temperature. When hanging up the laundry, they see that the shirts are still full of stains. Georges says to Yves: "This washing machine does an impeccable job." Having hung up the laundry, they go to drink a beer in the kitchen. Question: In your opinion, do Georges and Yves go to a pub?
Literal	Georges fait la lessive avec son colocataire Yves. Yves fait remarquer que Georges a beaucoup de chemises à laver. Ils passent un peu de temps à chercher le bon programme et la bonne température. Au moment d'étendre le linge, ils voient que les chemises n'ont plus aucune tache. Georges dit à Yves : « Cette machine fait un lavage impeccable. » Après avoir étendu le linge, ils vont boire une bière dans la cuisine. Question : A votre avis, est-ce que Georges et Yves vont dans un pub ?	Georges is doing the laundry with his roommate Yves. Yves points out that Georges has a lot of shirts to wash. They spend some time looking for the right program and the right temperature. When hanging up the laundry, they see that the shirts no longer have any stains. Georges says to Yves: "This washing machine does an impeccable job." Having hung up the laundry, they go to drink a beer in the kitchen. Question: In your opinion, do Georges and Yves go to a pub?
Ironic	Richard téléphone à son colocataire Gérald pour qu'il enregistre la finale de handball. Gérald est d'accord et réalise l'enregistrement. Dès que l'enregistrement est terminé, il le prépare pour Richard. Une fois rentré du travail, Richard le regarde avec Gérald. La qualité d'image est très mauvaise et Richard dit à Gérald « Ce lecteur fonctionne à la perfection. » Ils passèrent alors la soirée à regarder le match en mangeant des chips. Question : A votre avis, est-ce que Richard regarde un match enregistré ?	Richard calls his roommate Gerald to ask him to record the handball final. Gerald agrees to do it and he makes the recording. Once the recording is complete, he prepares it for Richard. After getting home from work, Richard watches it with Gerald. The picture quality is very poor and Richard says to Gerald: "This player works perfectly." They spend the evening watching the game and eating chips. Question: In your opinion, does Richard watch a match that was taped?
Literal	Richard téléphone à son colocataire Gérald pour qu'il enregistre la finale de handball. Gérald est d'accord et réalise l'enregistrement. Dès que l'enregistrement est terminé, il le prépare pour Richard. Une fois rentré du travail, Richard le regarde avec Gérald. La qualité de l'image est vraiment parfaite et Richard dit à Gérald : « Ce lecteur fonctionne à la perfection. » Ils passèrent alors la soirée à regarder le match en mangeant des chips. Question : A votre avis, est-ce que Richard regarde un match enregistré ?	Richard calls his roommate Gerald to ask him to record the handball final. Gerald agrees to do it and he makes the recording. Once the recording is complete, he prepares it for Richard. After getting home from work, Richard watches it with Gerald. The picture quality is really perfect and Richard says to Gerald: "This player works perfectly." They spend the evening watching the game and eating chips. Question: In your opinion, does Richard watch a match that was taped?

	French	English
Ironic	Eve et sa nouvelle colocataire Aline réfléchissent à ce qu'elles vont faire ce soir. Eve veut aller en boîte de nuit. Aline accepte et se prépare. En arrivant, elles constatent que la musique est démodée et que la piste est déserte. Aline se tourne vers Eve et dit : <u>« Ce soir la discothèque est palpitante. »</u> Cela faisait très longtemps que ni l'une ni l'autre n'étaient allées en boîte. Question : A votre avis, est-ce qu'Eve et Aline vont souvent en boîte ?	Eve and her new roommate Aline are thinking about what they will do tonight. Eve wants to go to a nightclub. Aline accepts and she gets ready. Upon arriving, they find that the music is outdated and that the dance floor is empty. Aline turns to Eve and says: <u>"This club is thrilling tonight."</u> It had been a long time since either of them had gone to a club. Question: In your opinion, do Eve and Aline go to clubs often?
Literal	Eve et sa nouvelle colocataire Aline réfléchissent à ce qu'elles vont faire ce soir. Eve veut aller en boîte de nuit. Aline accepte et se prépare. En arrivant, elles constatent que la musique est géniale et que la piste est pleine. Aline se tourne vers Eve et dit : <u>« Ce soir la discothèque est palpitante. »</u> Cela faisait très longtemps que ni l'une ni l'autre n'étaient allées en boîte. Question : A votre avis, est-ce qu'Eve et Aline vont souvent en boîte ?	Eve and her new roommate Aline are thinking about what they will do tonight. Eve wants to go to a nightclub. Aline accepts and she gets ready. Upon arriving, they find that the music is great and the dance floor is full of people. Aline turns to Eve and says: <u>"This club is thrilling tonight."</u> It had been a long time since either of them had gone to a club. Question: In your opinion, do Eve and Aline go to clubs often?
Ironic	Justine aide sa voisine Aurore à déménager. Justine transporte un gros carton. Aurore la remercie vivement car c'est le dernier carton. En fait, le carton est tellement lourd que Justine n'arrive pas à le soulever. Elle dit à Aurore : <u>« J'ai l'impression que ce carton est vide. »</u> Le soir Aurore fait un dîner pour sa crémaillère avec ses amis. Question : A votre avis, est-ce que Justine aide Aurore à déménager ?	Justine is helping her neighbor Aurore move. Justine is carrying a large cardboard box. Aurore thanks her profusely because it is the last one. Actually, the box is so heavy that Justine cannot manage to lift it. She says to Aurore: <u>"It's as though the box were empty."</u> That evening, Aurore cooks dinner to celebrate her housewarming with her friends. Question: In your opinion, does Justine help Aurora move?
Literal	Justine aide sa voisine Aurore à déménager. Justine transporte un gros carton. Aurore la remercie vivement car c'est le dernier carton. Le carton est très léger et Justine parvient à le descendre rapidement. Elle dit à Aurore : <u>« J'ai l'impression que ce carton est vide. »</u> Le soir Aurore fait un dîner pour sa crémaillère avec ses amis. Question : A votre avis, est-ce que Justine aide Aurore à déménager ?	Justine is helping her neighbor Aurore move. Justine is carrying a large cardboard box. Aurore thanks her profusely because it is the last one. The cardboard box is very light and Justine manages to bring it down quickly. She says to Aurore: <u>"It's as though the box were empty."</u> That evening, Aurore cooks dinner to celebrate her housewarming with her friends. Question: In your opinion, does Justine help Aurora move?

Appendix B: Decoys

French (as presented)	English translation
Matéo déménage et doit déplacer un miroir lourd et très fragile. Il demande à Paul de l'aider. Paul est disponible tout de suite. A peine a-t-il soulevé le miroir que ce dernier se brise en mille morceaux. Matéo dit à Paul: <u>« On a fait une grosse bêtise. »</u> Quelques jours plus tard, Matéo fête son emménagement avec des amis. Question: A votre avis, est-ce que Mateo et Paul ont déménagé le miroir sans problème?	Matéo is relocating and has to move a very fragile and heavy mirror. He asks Paul for help. Paul makes himself available immediately. As soon as Paul lifts the mirror it breaks into a thousand pieces. Mateo says to Paul: <u>"We have made a big mistake."</u> A few days later, Mateo celebrates his move with his friends. Question: In your opinion, do Matéo and Damien move the mirror without problems?
Damien et Myriam vont faire les soldes. Damien n'a pas une idée trop claire sur ce qu'il veut acheter. Myriam lui propose de tester différents magasins. En ressortant d'une cabine d'essayage, Damien est vêtu d'une	Damien and Myriam go shopping for clothes on sale. Damien doesn't have a very clear idea of what he wants to buy. Myriam suggests that he try different stores.

manière très extravagante. En voyant le résultat, il dit à Myriam : « Je n'aime pas ces habits. » Ils décident alors de passer au prochain magasin. Question: A votre avis, est-ce que Damien et Myriam vont chercher des nouveaux vêtements ?	Damien comes out of one dressing room clothed in a very extravagant manner. Seeing himself in the mirror, he says to Myriam: "I don't like these clothes." They decide to move on to the next store. Question: In your opinion, do you think Damien and Myriam are shopping for new clothes?
Emma organise une surprise pour l'anniversaire d'une copine. Elle demande à Romain de l'aider. Romain est tout à fait d'accord. Malheureusement, la copine en question découvre le secret une semaine avant la fête. Emma dit à Romain : « L'effet de surprise est complètement raté. » Cependant la fête a été un succès. Question: A votre avis, est-ce qu'Emma cherche à organiser une surprise ?	Emma organizes a surprise party for a (female) friend's birthday. She asks Romain to help her. Romain agrees. Unfortunately, the friend in question discovers the secret a week before the party. Emma says to Romain: "The surprise is completely ruined." Nonetheless the party was a success. Question: In your opinion, do you think Emma tried to organize a surprise party?
Hugo travaille dans une usine de voiture au service de la sécurité routière. Joël vient le voir le jour des crashs tests. Hugo lui montre les procédures pour les tests. Ils réalisent alors un test avec un mannequin assis dans la voiture. Le mannequin est complètement détruit à la fin du test et Hugo dit à Joël : « Il y a sûrement eu une erreur quelque part. » Hugo pense alors aux études faites avec les nouvelles normes de sécurité. Question : A votre avis, est-ce que le mannequin est intact après le test ?	Hugo works at a car factory in the department of road safety. Joel comes to see him on the day of the crash tests. Hugo shows him the testing procedures. They run a test with a dummy seated in the car. The dummy is completely destroyed at the end of the test and Hugo says to Joel: "There must have been an error somewhere." Then Hugo remembers the studies done using the new safety standards. Question: In your opinion, is the dummy intact after the test?
Martin part manger dans son restaurant habituel. Il commande un steak. Peu après avoir commandé, le serveur lui apporte son assiette. Le steak n'est pas du tout salé. Quand le serveur repasse, Martin lui dit : « Pouvez-vous m'apporter du sel ? » Après le plat principal, il commande un dessert délicieux. Question : A votre avis, est-ce que Martin va souvent dans ce restaurant ?	Martin goes toeat at his usual restaurant. He orders a steak. Shortly after he orders, the waiter brings him his dish. The steak is not at all salted. When the server returns, Martin says: "Could you bring me the salt?" After the main course he orders a delicious dessert. Question: In your opinion, does Martin go to this restaurant often?
L'imprimante de l'entreprise de Maxence était toujours en panne. Il demanda à sa secrétaire Nelly de la changer. Nelly commanda un nouveau modèle. La nouvelle imprimante se révéla être très compliquée à utiliser. Quand Maxence parvient par s'en servir, il dit : « Au moins elle imprime très bien. » Des clients arrivèrent et Maxence partit les accueillir. Question : A votre avis, est-ce que la nouvelle imprimante est facile à utiliser ?	The printer at Maxence's business is always broken. He asked his secretary Nelly to change it. Nelly orders a new model. The new printer turns out to be very complicated to use. When Maxence manages to use it, he says: "At least it prints very well." Some clients are arriving and Maxence goes to greet them. Question: In your opinion, is the new printer easy to use?

Table 1. A summary of neuroimaging studies to date on irony and the extent to which they implicate ToM activity as based on generally accepted description of the neural ToM.

Study	Theory of Mind Network				Comments
	rTPJ	*lTPJ*	*MPFC*	*PC*	
Eviatar & Just, 2006	None reported	None reported	Yes	None reported	MPFC activation in all conditions
Wang et al., 2006	None reported	None reported	Yes	None reported	
Uchiyama et al, 2006	None reported	None reported	Yes	None reported	No contrast: *Irony > Literal*
Wakusawa et al., 2007	None reported	None reported	Yes	None reported	
Rapp et al, 2010	None reported	None reported	Yes	None reported	
Shibata et al., 2010	None reported	None reported	Yes	Yes	
Uchiyama et al., 2011	None reported	None reported	Yes	None reported	No contrast: *Irony > Literal*

Table 2. An example from the *Ironic* condition, its *Literal* control as well as a decoy (translated from French).

Condition	Example
Ironic	Cynthia and Léa sing together in the same opera.
	On the night of the premiere they meet at the theatre.
	The show begins exactly on time.
	During their performance both ladies sing off key.
	After the show, Cynthia says to Léa:
	"Tonight we gave a superb performance."
	As they take off their make-up they continue to discuss the show.
	Question: Do you think that the performance was in the morning?
Literal	Cynthia and Léa sing together in the same opera.
	On the night of the premiere they meet at the theatre.
	The show begins exactly on time.
	Both ladies sing beautifully and receive a rapturous round of applause.
	After the show, Cynthia says to Léa:
	"Tonight we gave a superb performance"
	As they take off their make-up they continue to discuss the show.
	Question: Do you think that the performance was in the morning?
Decoy	Mateo is relocating and has to move a very fragile and heavy mirror.
	He asks Paul for help.
	Paul makes himself available immediately.
	As soon as Paul lifts the mirror it breaks into a thousand pieces.
	Mateo says to Paul:
	"We have made a big mistake."
	A few days later, Mateo celebrates his move with his friends.
	Question: In your opinion, do Mateo and Damien move the mirror without problems?

Table 3. Brain areas activated in the contrast *Ironic* condition > *Literal* condition. Notes. L., left; R. right; ~ BA, approximate Brodmann's area; Cluster-wise FDR (false discovery rate) corrected *p* < .05 and voxel-wise uncorrected *p* < .001.

Anatomical locations	~ BA	No. of Voxels in Clusters	Talairach coordinates			z score
			x	y	z	
			6	56	29	5.18
RL medial prefrontal cortex	6/8/9	541	-6	39	40	5.11
			3	49	12	4.63
L inferior frontal gyrus	45/46/47	303	-42	25	-3	4.34
L inferior parietal lobule	40		-54	-62	48	4.70
L temporal parietal junction	40	101	-56	-58	34	3.28
L insula	47	25	-36	16	-13	4.93
R dorsolateral prefrontal cortex	8	61	42	13	44	4.62
R middle temporal gyrus	21	113	62	-22	-8	4.46
R temporal parietal junction	40	126	62	-55	29	4.41
R inferior frontal gyrus	45/46/47	196	59	-22	16	4.25

Figure legends

Figure 1. Experimental procedure.

Figure 2. Contrast: *Ironic* condition > *Literal* condition. Cluster-wise FDR corrected p < .05 and voxel-wise uncorrected $p < .001$. *The activation in the PC is significant only if we use cluster-wise FDR corrected $p < .05$ and voxel-wise uncorrected $p < .005$.

Figure 3. ROIs of the four main areas of the ToM network. The coordinates come from a review of the neuroscientific literature on Theory of Mind performed by van Overwalle and Baetens (2009). MPFC (x = 0 y = 50 z = 20), rTPJ (x = 50 y = -55 z = 25), lTPJ (x = -50 y = -55 z = 25), PC (x = 0 y = -60 z = 40). There was significantly greater activity in the *Ironic* than *Literal* condition in all four regions. Specifically, for the rTPJ: t(19) = 2.42, $p = .026$, for the MPFC: t(19) = 2.92, $p = .009$, and for the PC: t(19) = 2.31, $p = .032$), whereas the lTPJ showed an activation increase that was marginally significant (t(19) = 1.98, $p = .062$). The graphs report the percentage signal change for the target sentence of the two critical conditions, for each ROI.

Figure 4. Psychophysiological interaction (PPI) analysis. It shows the increase of functional connectivity between the ventral part of the MPFC (seed; x = 0 y = 55 z = 6) and the left IFG (peak: x = -50 y = 31 z = 0) for the contrast *Ironic > Literal*. There is also an increase of functional connectivity between the MPFC and the right IFG (peak: x = 53 y = 18 z = 25 x = -50 y = 31 z = 0) for the same contrast.

Figure 1.

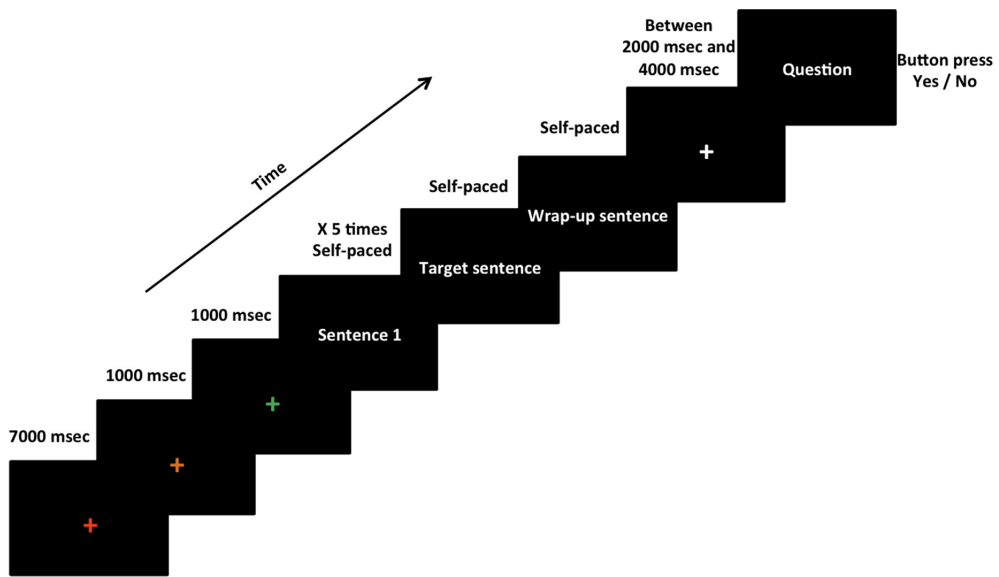ToM-related regions

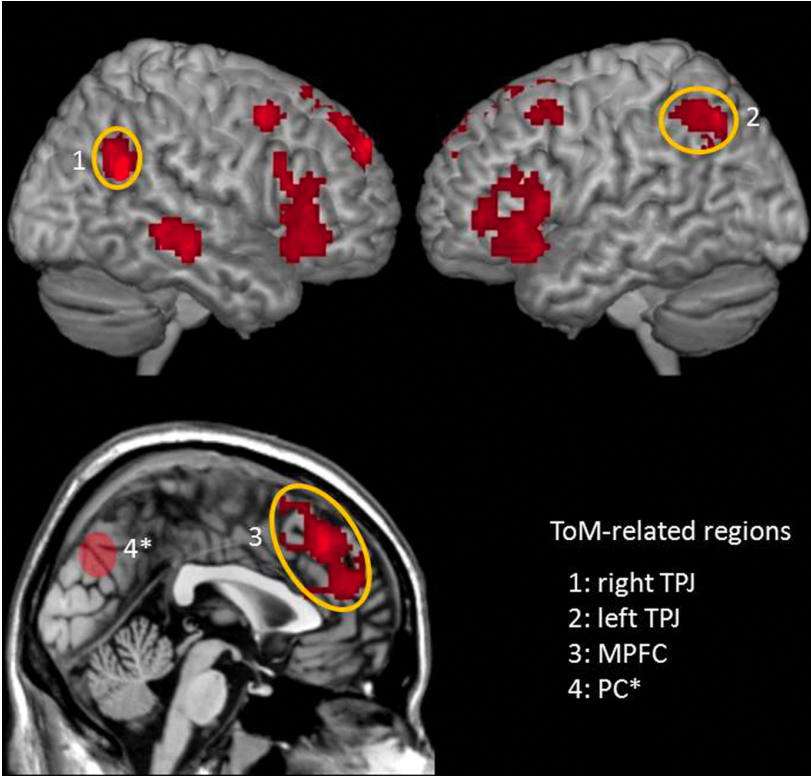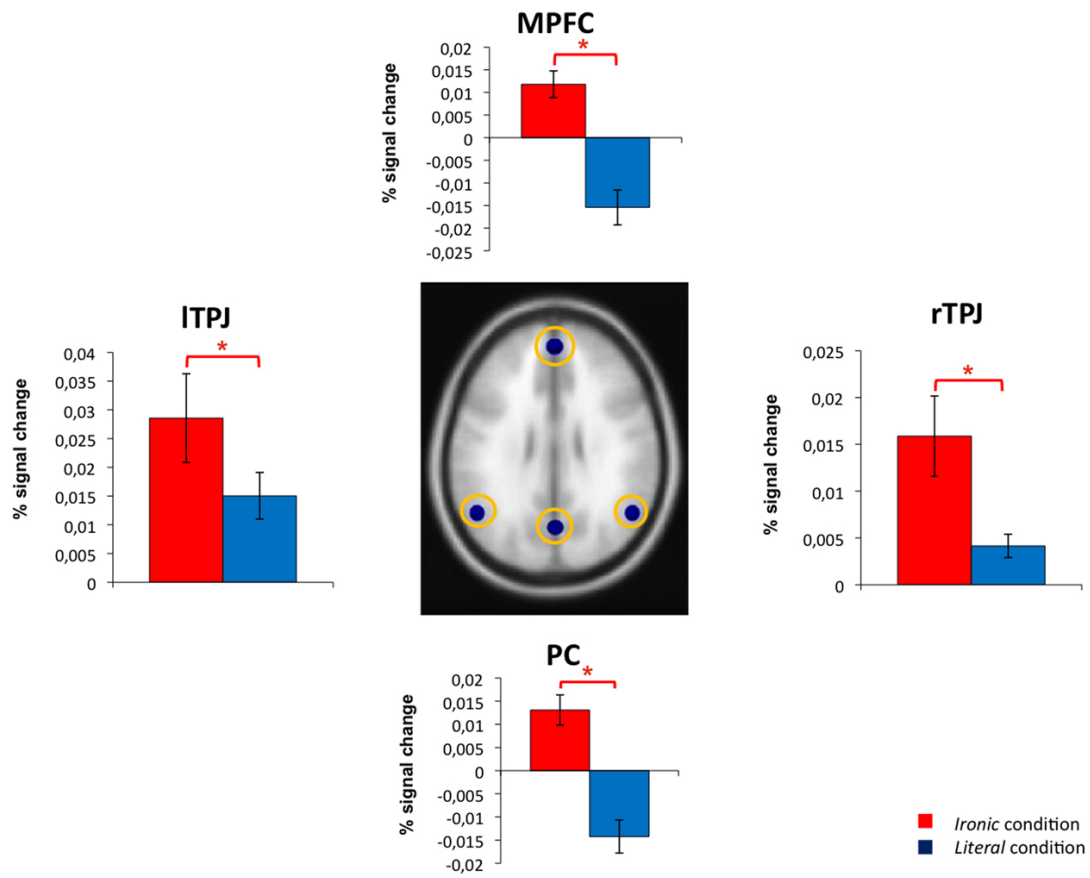1: right TPJ
2: left TPJ
3: MPFC
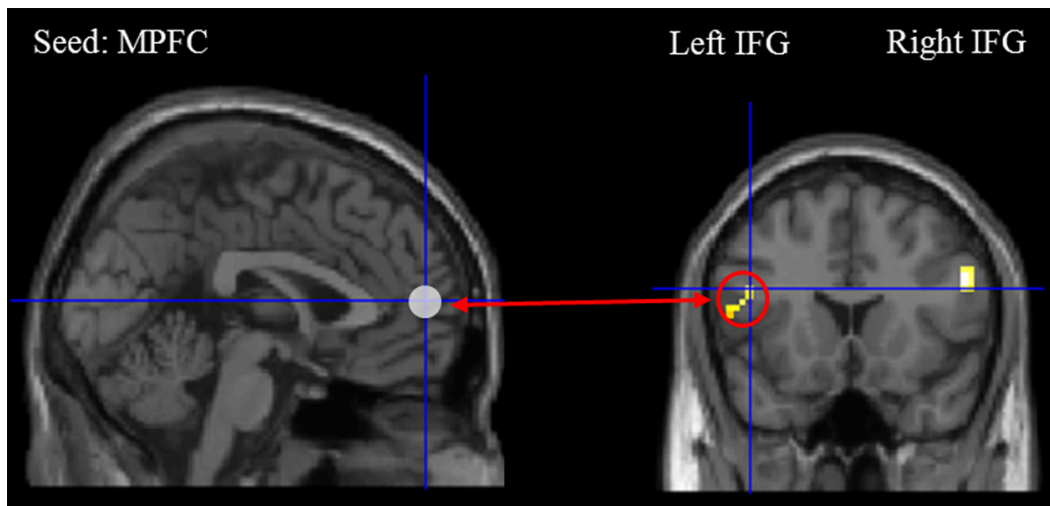4: PC*

Figure 2.

Figure 3.

Figure 4.

7. The time course of irony processing

The fMRI study revealed evidence showing that the Theory of Mind network is critically engaged during irony processing. As much as we are indebted to fMRI methodology, its temporal resolution is weak and does not allow us to investigate the allocation of different cognitive resources during the on-line processing of an ironic utterance. In this case the literature on the time course of ToM cannot help us because there are only a couple of ERP studies on Theory of Mind. Below, I outline the main results of the few ERP experiments on ToM as well as those coming from the previous ERP investigations on irony.

7.1 <u>Temporal dynamic of Theory of Mind processing</u>

fMRI and lesion studies have been very helpful in identifying the ToM network, but they do not allow researchers to investigate the temporal dynamic of ToM processing. Whereas fMRI has a spatial resolution of a couple of millimeters, it has a temporal resolution of just a couple of seconds. That said, the recording of an EEG signal has been used sparingly to focus on ToM (e.g., Geangu, Gibson, Kaduk, & Reid, 2012; Liu, Sabbagh, Gehring, & Wellman, 2004; J Meinhardt, Sodian, & Thoermer, 2011; Jörg Meinhardt, Kühn-Popp, Sommer, & Sodian, 2012; Sabbagh & Taylor, 2000). A critical challenge for EEG studies on ToM is that the signal recording is concentrated on a time window that lasts just 1 to 1½ seconds. While the phenomenon under investigation should strongly affect cognitive systems during this very short period, ToM is probably a long-lasting process that is hard to grasp in such a brief interval. Consider, for example, a standard false belief task: When can one suppose that the subject starts to interpret the character's beliefs? Sabbagh & Taylor (2000) introduced a question at the end of each story in order to have a precise trigger. The question was about the character's beliefs (e.g., "According to the character, where is the [Object A / Object B]?"). Both this study and another study based on a similar design (Liu et al., 2004) linked a left-frontal lateralized slow wave with ToM processing. However, the presence of the question turns the paradigm into an explicit reasoning task leading to two risks. First, the conscious reasoning on others' thoughts may or may not engage precisely the same processes as an on-line mindreading activity. Second, the subject may have already interpreted the character's thoughts during the presentation of the story and so, when recording ERPs only on the question, one may be grasping phenomena other than ToM. Therefore, the investigation of

the temporal dynamics of ToM processing would be greatly improved by the development of a paradigm that triggers mindreading activity at a reasonably precise moment.

7.2   Previous EEG studies on irony

As far as we know there are already 5 EEG studies on irony (Amenta & Balconi, 2008; Cornejol et al., 2007; Katz et al., 2004; Regel, Coulson, & Gunter, 2010; Regel et al., 2011) and the most recurrent result among the different designs is that irony processing elicits a larger P600/late positivity than the processing of its cohort literal sentences. Unfortunately, the functional interpretation of the P600 is difficult because a wide range of linguistic phenomena elicit enhancement of this component.

The P600 was originally associated with syntactic violations (e.g., violation of phrase structure, Friederici & Meyer, 2004) but enhancement of the P600 has been reported for a broad range of linguistic expressions such as semantic reversal anomalies (e.g., "The cat that fled from the mice.", van Herten et al., 2005) and semantic anomalies (semantic illusions) in a discourse context (e.g., "How many animals of each sort did Moses put on the Ark?", see for example Nieuwland & Berkum, 2005). The P600/Late positivity effect is also one of the most recurring outcomes of studies on pragmatic phenomena. For example, indirect requests (see for example, Coulson & Lovett, 2010), metaphors (see e.g., De Grauwe et al., 2010) and ironies (see for example Regel et al., 2011) elicit an increase of the P600 component. However, considering the broad range of linguistic phenomena that affect the P600, it is hard to come up with a functional interpretation of the P600 in the pragmatic domain. A possible *fil rouge* of pragmatic phenomena is that they usually involve a mismatch between the literal meaning and the message conveyed by the utterance, and this inconsistency matches with one of the major interpretations of the P600 effect that would link the increase of the P600 to the increase in cognitive load during the integration processes (for reviews see, for example, Bornkessel-schlesewsky & Schlesewsky, 2008; Hagoort, Brown, & Osterhout, 1999). However, this interpretation is not yet established in the literature.

The ERP studies on irony have not pushed the debate in a new direction; on the contrary, they have remained focused on the psycholinguistic accounts of irony. In the most recent one (Regel et al., 2011), for example, authors aimed to determine the extent to which their results verify the predictions of the Standard Pragmatic Model, the Direct Access View and the Graded Salience Hypothesis. The data revealed just a partial match with both the

Standard Pragmatic Model and the Graded Salience Hypothesis and so Regel et al. called for a revision of the psycholinguistic models of figurative language comprehension.

As we have argued extensively, we believe that the psycholinguistic debate has veered from the most interesting (and relevant) path with respect to irony processing. In our EEG study we try to stay focused on the allocation of different cognitive resources during pragmatic inference making more than on the hypotheses about the priority of the literal meaning of an ironic utterance.

8. One ironic second

When we talked about the fMRI activation of the Theory of Mind network during irony processing the data were averaged on an interval of not less than 2 seconds; this is a well known limit of fMRI measures and so one could say that our study has provided a picture of irony processing that only arrives at the end of the dance. In order to investigate what happens *while* an irony is understood we decided to employ here the same paradigm of the second behavioral experiment (the one we used in the fMRI study). In this way, we are in the position to obtain information on irony processing in "real time". While EEG and the study of ERP components have been helpful for clarifying the functional interpretation of linguistic phenomena, such analyses do not allow us to infer what kind of cognitive resources are allocated at a given moment. In order to partially bypass this issue we decided to analyze not only the ERPs but brain oscillations as well. These put us in a better position to understand the ongoing processes based on the increase and decrease of power in the different frequency bands (see for example, Bastiaansen et al., 2012).

We expected to find an increase of the P600 during the *Ironic* condition when it is compared to the *Literal* one. In addition, we concentrated our attention on three frequencies bands: theta (4 – 7 Hz), alpha (8 – 12 Hz) and gamma (> 30 Hz), in order to investigate the retrieval, encoding and integration opeations in irony processing.

What are you doing while I am speaking to you?
The allocation of cognitive resources during irony processing

Nicola Spotorno[a], Anne Cheylus[a], Jean-Baptiste Van Der Henst[a], Ira A. Noveck[a,b]

[a] Laboratory L2C2, CNRS Université Lyon 1 FRE3406, 69675 Bron CEDEX, France

[b] Centre de recherche français à Jérusalem (CRFJ), MAEE-CNRS (UMIFRE 7) B.P. 547, 3 rue Shimshon, Baka, 91004 Jérusalem, Israel

Corresponding author:

Nicola Spotorno, L2C2 CNRS-Université Lyon 1, Institut des Sciences Cognitives, 67 Bd Pinel - 69675 Bron CEDEX, France.

Phone: +33 (0)4 37 91 12 71

E-mail address: nicola.spotorno@isc.cnrs.fr

Abstract

From the must banal expressions like "The cat is on the mat" to the most poetical and cryptic ones the knowledge of the words meaning and of grammatical rules do not ensure one to grasp the message the speaker wants to communicate. In order to understand the complete meaning of an utterance one has to integrate the linguistic code with contextual information and the so-called pragmatic inferences allow one to catch what the speaker has in mind. The actual computation of a pragmatic inference engages several cognitive resources that integrate different streams of information in a single representation of the utterance meaning.

The present work has been designed to investigate the allocation of different cognitive resources during pragmatic inference making by the recording of the EEG, but we needed the right test bed. We chose irony because the message the speaker wants to communicate with an ironic remark is completely beyond the linguistic code. The results will show that irony processing elicits an enhancement of the P600 component along with an increase of power in the theta bad that can be related with the increase of load in the short-term memory in order to encode new information into the episodic memory. In addition, the decrease of power on the left parietal areas of the scalp in the alpha band can be related to the recruitment of attentive resources during irony processing, while the increase of power in the gamma band may reveal the integration between different cognitive resources, namely language and Theory of Mind, since the first steps of the comprehension of an ironic utterance. As long as we know, this is the first EEG study that employs the time frequencies analysis in the investigation of a pragmatic phenomenon.

1. Introduction

Imagine two fishermen who throughout their one-day-long fishing trip do not catch a fish. When one fisherman tells the other (1), he is clearly being ironic.

(1) "This day has been really productive!"

In order to understand the ironic remark the second fisherman has to go beyond the linguistic code and make pragmatic inferences (e.g., Grice, 1975; Sperber & Wilson, 1986), which are processes that integrate the linguistic stimulus with contextual information. Irony is a clear case of a pragmatic phenomenon because what the speaker says is not at all what the speaker wants to communicate, and so pragmatic inferences are crucial for one to grasp the meaning of the ironic remark.

An interpretation is encoded in the linguistic stimulus when it is stipulated by the grammar, but since Grice's William James Lectures (reprinted in Grice, 1989) a sustained and largely successful attack on unreflective appeals to encoding, the borderline between linguistic communication and linguistic encoding has been a major focus of pragmatic research. Now, one of the main assumptions of Pragmatics is that the linguistic code almost always underdetermines the complete meaning of an utterance. In addition, daily experience supports the importance of pragmatic inferences in language comprehension: from simple deictic expressions like "This is an apple." to more cryptic ones such as "No man is an island." the simple juxtaposition of word meanings does not provide the intended message of an utterance.

According to both Gricean and post-Gricean perspectives, the critical skill that allows one to fill in the gap between the linguistic code and utterance interpretation is the ability to interpret the speaker's intentions. From the point of view of Pragmatics the speaker is not a black box that provides only strings of code as output; on the contrary, in every conversation we view our interlocutors as intentional agents who are moved by beliefs and desires and who try to share, impose or suggest parts of their inner world. To make it more concrete, imagine that Mary has had hard time at her job. As usual she fought with her boss and he threatened to fire her if she does not stop contradicting him. When she comes back home her husband Paul asks about her day and she sadly replies (2):

(2) "As usual."

The linguistic code dramatically underdetermines the meaning of (2). Mary probably wants to communicate that she always has a hard time at her job and that that given day is not an exception. Paul should be able to get the message because he pays attention to every aspect of Mary's behavior that seems relevant to him (e.g., her sad tone of voice). In addition, he shares with Mary a common ground about her current situation at the job and he knows that she knows that he knows that she is currently having a hard time at her job. It is evident that the comprehension process goes further than the decoding of (2) and that Paul will be able to completely appreciate the meaning of (2) only if he has access to Mary's mental state.

From a psychological perspective, the ability to interpret others' intentions is usually called Theory of Mind (ToM).

> "In saying that an individual has a theory of mind, we mean that the individual imputes mental states to himself and to others (…); A system of inferences of this kind is properly viewed as a theory, first, because such states are not directly observable, and second, because the system can be used to make predictions, specifically about the behavior of other organisms." (Premack & Woodruff, 1978, 515)

ToM is supposed to guide our behavior in every aspect of our interactions with other people. In particular, several studies on children have shown that ToM and pragmatic abilities have similar developmental trajectories (e.g., Bloom, 2000; Filippova & Astington, 2008; Happé, 1993; Langdon, Davies, & Coltheart, 2002). For example, at around the age of 4 children are able to pass the false belief test (which tests children's ability to take into account the beliefs of someone else when those beliefs diverge from their own; Wimmer & Perner, 1983) and, at around the same age, they start to understand metaphors (which require pragmatic inferences to be understood). Around the age of 7 children are able to pass second-order false belief tests (i.e., a task in which children have to take into account the false beliefs of a character about the beliefs of another character) and around 7 - 9 years old they start to correctly interpret ironic statements (e.g., Filippova & Astington, 2008). In addition, studies on Autism spectrum disorders (ASD), have revealed that ASD affects both ToM and pragmatic abilities (e.g., Happé, 1993).

Over the last 20 years, cognitive neuroscience has investigated the neural basis of both language and ToM. As a result, there is now a general agreement about the candidate regions for both language and ToM networks (see Ferstl, Neumann, Bogler, & Von Cramon, 2008; Hagoort, 2005) for a review on the language network and Van Overwalle & Baetens, 2009 for a meta-analysis on ToM literature). However, the investigation of the interactions between linguistic and ToM abilities during the on-line comprehension of an utterance is still only starting.

A recent fMRI study on irony processing by our group (Spotorno et al., 2012) has revealed that the entire, bilateral ToM network is more active during the processing of an ironic utterance than during the processing of the same sentence understood literally. In addition, Spotorno and colleagues conducted a Psychophysiological Interaction Analysis (PPI), which allows one to investigate the functional connectivity among different regions. The result of the PPI showed an increase of functional connectivity between the medial prefrontal cortex, which is considered one of the core regions of the ToM network, and the left inferior frontal gyrus, which is involved in integration processes in the linguistic domain. That is, the PPI analysis crucially revealed an interaction between language and ToM networks during irony processing.

As is well known, fMRI analyses are not ideal with respect to temporal resolution which prevents the drawing of strong conclusions about pragmatic inference making *while* they are computed.   However, in order to bypass this issue, several researchers have conducted ERP studies on pragmatic phenomena (e.g., Grauwe, Swain, & Holcomb, 2010; Nieuwland, Ditman, & Kuperberg, 2010; Noveck & Posada, 2003; Regel, Coulson, & Gunter, 2010; Regel, Gunter, & Friederici, 2011). This is what we turn to here.

1.1 Pragmatic inferences and ERP components

The N400 and the P600 are the ERP components that have received the most attention in the investigation of pragmatic inferences. The N400 is an index of semantic/world knowledge violations and, historically, it has been associated with strong violations at the sentence level such as: "He spread the warm bread with socks" (Kutas & Hillyard, 1980). However, it has been shown that the N400 can more generally reflect inconsistencies with respect to the context (see, for example Nieuwland & Van Berkum, 2006). The P600 was originally considered to be related with syntactic violations (e.g., violation of phrase structure, Friederici & Meyer, 2004) but it has also been associated with a broad range of linguistic expressions

such as semantic reversal anomaly (e.g., "The cat that fled from the mice", Herten, Kolk, & Chwilla, 2005) and semantic anomalies (semantic illusions) in a discourse context (e.g., "How many animals of each sort did Moses put on the Ark?" (see e.g., Nieuwland & Berkum, 2005). The P600 / Late positivity effect is also one of the most recurring outcomes from studies on pragmatic phenomena. For example, indirect requests (see for example, Coulson & Lovett, 2010), metaphors (see e.g., Grauwe et al., 2010) and ironies (see e.g., Regel et al., 2011) elicit an increase of the P600 component. However, considering the broad range of linguistic phenomena that affect the P600, it is hard to come up with a functional interpretation of the P600 in the pragmatic domain. Pragmatic phenomena usually involve a mismatch between the literal meaning of an utterance and the message it conveys (e.g., when one says "John is a shark" they do not mean that John is a real shark; likewise, when (1) is used ironically it is meant to indicate that the day was very unproductive) and this inconsistency matches with one of the major interpretations of the P600 effect, which links an increase in the P600 to an increase in cognitive load during the integration processes (for reviews see, for example, Bornkessel-schlesewsky & Schlesewsky, 2008; Hagoort, Brown, & Osterhout, 1999).

The association of the P600 with integration processes is already informative, but analyses through another technique, brain oscillations, may help to disentangle the different cognitive resources that are involved in pragmatic inference making. Unfortunately, the application of this technique to language comprehension is still in its nascence, and so the literature does not provide a stable framework against which one can interpret new data. Nevertheless, researchers such as Bastiaansen (e.g., Bastiaansen, Mazaheri, & Jensen, 2012), Klimesch (e.g., Klimesch, 1999) and Weiss (e.g., Weiss et al., 2005) have provided a first sketch of the oscillatory dynamics of the brain during language comprehension.

1.2 Brain oscillations

Time frequency analysis (TF) allows one to investigate the patterns of synchronization and desynchronization of neuronal activity related to the coupling and uncoupling of functional networks in the brain (see e.g., Pfurtscheller & Lopes da Silva, 1999; W. Singer, 1993; Wolf Singer, 1999; Varela & Lachaux, 2001). Elements pertaining to the same functional network are identifiable as such by virtue of the fact that they fire synchronously at a given frequency and then synchronous oscillations in a wide range of frequencies are considered to play a crucial role in linking areas that are part of one and the same functional network. One of the

main features of the oscillations is that they are ongoing phenomena that occur even in the absence of any experimental task. As a result, the phase of the oscillation at the moment of the event is variable. This is the reason why such event-related changes in oscillatory EEG activity are considered non-phase-locked responses (while, on the contrary, ERPs are phase-locked responses). However, the experimental stimuli modulate the oscillatory activity and so event-related but non-phase-related responses may be meaningfully related to the event in question like, for example, language comprehension. Considering that synchronization/desynchronization reflects the coupling/decoupling of neural networks, it can be argued that these oscillatory EEG responses provide a window into the functional network dynamics of the brain (see for example Bastiaansen et al., 2012).

1.2.1   Brain oscillations in language comprehension

Different aspects of language comprehension have been associated with several different frequency bands, namely theta (4 – 7 Hz), alpha (8 – 12 Hz), lower beta (13 – 18 Hz) and gamma (above 30 Hz). Synchronization (in the following will be used to indicate also as increase of power in the frequency band) in the theta band has been associated with both the retrieval of lexical information and the encoding of new information into episodic memory. For example, Bastiaansen, van der Linden, Ter Keurs, Dijkstra, & Hagoort, 2005) found an increase of power in the theta band in the left parietal areas of the scalp for open class words (e.g., adjectives and nouns), which are the words that convey the majority of semantic information of a sentence, in comparison to closed-class words (e.g., articles and prepositions), which are less charged in terms of semantic information but which connect the different components of a sentence. In addition, Bastiaansen and colleagues (Bastiaansen, Oostenveld, Jensen, & Hagoort, 2008) have shown that the scalp distribution of the increase of power in theta frequency varies as function of the kind of semantic information conveyed by the word. During a lexical decision task they presented both nouns with visual semantic properties, such as colors or shapes, and nouns with auditory semantic properties (ones that refer to sounds). Words with auditory semantic properties have shown a larger increase of power in the theta band in the scalp areas overlying the left auditory cortex than in the areas overlying the left visual cortex, while words with visual semantic properties elicited the opposite pattern. Considering that the only difference between the two sets of words lay in their semantic properties, the results support the hypothesis that neuronal synchronization in the theta frequency range is involved in the retrieval of lexical information. Another process

that elicits an increase of power in the theta band is the encoding of new information into episodic memory (see e.g., Klimesch, 1994, 1996). Klimesch (1996) tested this hypothesis by trying to isolate the pure encoding process from confounding variables such as attentional demands, task difficulty and cognitive load. Subjects performed a lexical decision task and they were not informed that they would be tested for memory retrieval of the words after the on-line session. The authors measured the theta power during the encoding process and they compared the synchronization in the theta band for the words that were retrieved after the session with that for the words that were not retrieved. The results revealed that the words that were remembered in the later recalling session elicited a larger increase of power in the theta band during the encoding stage with respect to the words that were not remembered. Klimesch and colleagues claimed that the design of their study allowed them to exclude differences with respect to mental load and experiment-related encoding strategies leading them to argue that the only difference between those words that were remembered over those that were not remembered was the establishment of a "memory trace".

In the context of language processing, the desynchronization in the alpha band has also been associated with memory retrieval operations (e.g., Klimesch, Doppelmayr, Pachinger, & Ripper, 1997; Klimesch, Doppelmayr, Schimke, & Ripper, 1997; Sterman, Kaiser, & Veigel, 1996). Klimesch and colleagues have especially noted that desynchronization in the upper alpha band (broadly 10 - 12 Hz) positively correlates with semantic long-term memory performance (for a review see Klimesch, 1999). For example, in a semantic judgment task, Klimesch et al. (Klimesch, Doppelmayr, Pachinger, et al., 1997) presented "concept – feature" pairs (e.g., "claws – eagle", "wings – banana") and participants had to judge whether the two words were semantically congruent. The results showed an increase of desynchronization in the upper alpha band for the concept words when compared with the feature words, in line with the authors' predictions that desynchronization in the alpha band correlates with semantic memory operations.

In a recent review of the time frequency literature by Bastiaansen et al. (2012) the authors emphasized that, beyond the encoding and retrieval operations, the unification of different streams of information is a critical stage in language processing. Bastiaansen and colleagues argued that beta and gamma bands reflect unification operations in syntactic and semantic domains respectively. In addition, in a study by Weiss et al., 2005 the lower beta frequency range (13- 18 Hz) was found to be sensitive to the syntactic complexity of the stimuli while several researchers found that gamma power correlates with semantic/world knowledge complexity (e.g., Hagoort, Hald, Bastiaansen, & Petersson, 2004; S. Weiss,

Rappelsberger, Schack, & Mueller, 2003; van Berkum, Zwitserlood, Bastiaansen, Brown, & Hagoort, 2004). For example, Hagoort et al. (2004) presented subjects with three versions of the same sentence differing only in one adjective (3a - c):

(3a)   "The Dutch trains are *yellow* and very crowded."
(3b)   "The Dutch trains are *white* and very crowded."
(3c)   "The Dutch trains are *sour* and very crowded."

It is common knowledge in the Netherlands that the Dutch trains are *yellow*, which makes (3b) incongruent with the world knowledge of Dutch subjects, despite the fact that a train can be *white*. On the contrary, (3c) carries a stronger/semantic violation, because *sour* usually refers to the taste of food. The results of this study showed an increase of power in the gamma band when the authors compared the correct utterances (e.g., 3a) with the incorrect but plausible ones (e.g., 3b), while the contrast between the correct utterances and the incorrect and implausible ones (e.g., 3c) did not show an increase of power in the gamma band. These data suggest that gamma oscillations reflect unification processes in the semantic domain and, interestingly, that the distinction between world knowledge and semantic violation can be detected by time frequency analysis but not by ERP analysis; both kinds of violations elicit an increase of the N400 component.

In addition, a study by van den Brink et al. (2010) has shown that the integration of "semantic" and "social" information into the linguistic context affects different frequency bands. In their study, lexical semantic violations elicited an increase in the theta range across all participants, but only individuals with an empathizing-driven cognitive style (as revealed by a psychological test) showed a larger increase of power in the gamma band in the presence of speaker's identity violations (e.g. when a sentence like "I cannot sleep without my teddy bear" is said with an adult-sounding voice).

This short and partial review of the literature suggests that the analysis of brain oscillations can provide interesting insights into language comprehension and that it can help disentangle the different cognitive resources that play a role in pragmatic inference making. However, the picture is still incomplete because the number of studies is still limited and because, as mentioned by Bastiaansen et al. (2012), "so far, experimental research into the oscillatory neuronal dynamics of unification operations have concentrated on semantic and syntactic unification only" (Bastiaansen et al. 2012, page 26). Language comprehension requires one to take into account pragmatic (and often phonological) information as well.

1.3 The present study

The present study has been designed to investigate the allocation of different cognitive resources during pragmatic inference making. In order to accomplish this aim we chose irony as a test bed because, in line with both theoretical positions (e.g., Clark & Gerrig, 1984; Grice, 1989; Wilson & Sperber, 2012) and experimental results (e.g., Spotorno et al., 2012), we claim that irony processing critically involves the integration of different cognitive resources, namely, language and ToM ability.

As far as we know there are already 5 EEG studies on irony (Amenta & Balconi, 2008; Cornejol et al., 2007; Katz et al., 2004; Regel et al., 2010, 2011) and the most recurrent result among the different designs is that irony processing elicits a larger P600/late positivity component when compared to the processing of literal sentences. As reported at the beginning, this result is in line with a general tendency reported in the ERP literature with respect to pragmatic phenomena, but the functional interpretation of the P600 is hard because a wide range of linguistic phenomena elicit enhancement of this component. In order to deeply investigate irony processing, we will conduct both the ERPs and the TF analysis. This is the first study that applies TF analysis to irony processing and, following the remark of Bastiaansen et al. (2012, p 26), it is also the first study that employs TF analysis in the investigation of integration operations during the on-line processing of a genuine pragmatic phenomenon.[5]

2. Materials and methods

2.1 Participants

Twenty healthy participants, who were students from the University of Lyon, participated in the study. All participants (whose mean age was 23) were native French speakers, were right-handed and reported to have normal vision and no history of mental illness. Our protocol was accepted by the local ethics committee and each participant gave informed consent prior to the experiment.

---

[5] In Hagoort et al., 2004 and Hald et al., 2006 the authors manipulated the consistency of the sentences with subjects' semantic and world knowledge as experimental variables. The integration of the linguistic code and world knowledge already figures into the domain of Pragmatics; our remark (i.e., "genuine") just underlines that we will study the comprehension process of correct and common expressions and not the processing of violations that are experimentally introduced.

2.2 Materials

Sixty story-frameworks were created (in French) that led to a target sentence that could be interpreted either as ironic or as literal as a function of a minor modification made to the prior context. In the *Ironic* condition, a target sentence, for example (4) "Tonight, we gave a superb performance" was preceded by a negative context (e.g., a terrible performance) whereas in the *Literal* condition the target sentence was preceded by a positive context (e.g., an impressive performance). Otherwise, the introductory sentences and the wrap-up sentences of any given framework were the same for both conditions. The structure of the stories was the same as the one employed in Spotorno et al. (2012). Each story included the following six features:

First, all stories were seven lines long, each having a maximum length of 91 characters (spaces included) in order to fit into one line on a screen. Second, the stories described an everyday situation and an exchange between two characters who know each other casually (i.e., we avoided situations that presumed close relationships among interlocutors). Third, the first three sentences introduced the two characters and the situation. Fourth, the fourth and fifth sentences described the development of the situation that can be either positive (in the literal version) or negative (in the ironic version). These were the only two lines that could potentially change with respect to condition. Changes were designed to be as minimal as possible while keeping the stories sensible. Fifth, the sixth line was designed to be the target sentence. The length of all target sentences was between 10 and 12 syllables. Crucially, the target sentence (line 6) is exactly the same in both *Ironic* and *Literal* conditions and just the last word (e.g., "productive" in 1) allows one to evaluate if the sentence is ironic or not. Finally, the seventh line was an ordinary wrap-up conclusion of the story that makes sense for both the *Ironic* and *Literal* conditions (for an example see Table 1).

We aimed to prevent negative contexts from being cues to the presence of ironic remarks introducing stories in which a negative context (e.g., a bad performance) leads to a plain, non-ironic utterance (e.g., "We will do better the next time."). We refer to these stories as *decoys*. We created 30 decoys and their structure was the same as the ironic stories (7 sentences in which a negative event occurs), except that the target sentence was banal. For example, the decoy story in Table 1 describes how one character drops a mirror, which leads the other character to remark "We have made a big mistake." Like in the *Ironic* and *Literal* conditions, the target sentence in the decoy stories is between 10 and 12 syllables. We designed also 30 positive fillers in which a positive context was followed by a positive remark.

Insert Table 1 about here

Each participant read 30 ironic stories, 30 literal stories, 30 decoys and 30 positive fillers. For each participant, the 60 critical (non-decoy) stimuli were extracted randomly from a pool of 60 frameworks that could each be the basis of either an ironic or literal target sentence. The 30 decoys and the 30 positive fillers remained the same for each participant.

To verify that the stimuli used here were perceived as intended, a rating study was conducted on the 90 stories (2 from each of the frameworks plus the 30 decoys) with 42 participants (22 women) whose ages ranged from 19 to 38 (with a mean of 26) and who did not participate in the EEG study. Participants were asked to read each story and rate the extent to which the target sentence was ironic on a scale from 1 (not at all ironic) to 5 (very ironic). Ironic target sentences were rated as highly ironic (mean of 4.3), while literal sentences and the banal lines from the decoy stories were rated as low on the ironic scale (1.3 both of them). Repeated measure ANOVAs showed significant differences between (i) the Ironic and Literal conditions and (ii) the Ironic condition and Decoys (both at $p < .001$, corrected for multiple comparisons using the Tukey method). The comparison between the Literal condition and the Decoys was not significant ($p = .1$)

A yes/no comprehension question followed one third of the items (regardless of whether it was a critical or filler item). The question was about a detail in the story that made no reference to the target sentence. The purpose of the comprehension question was to ensure that the participants were paying attention to the stories. (See Table 1 for an example of all conditions and questions and the Appendix for further examples.)

2.3 Procedure

Stimuli were prepared with Presentation 11.0 software (Neurobehavioral Systems, www.neurobs.com) and presented on a computer screen. Participants performed the experiment in four runs of 30 stories each. Each trial started with the presentation of a visual fixation mark (a white central cross). The participant read the stories line by line (i.e.,

sentence by sentence) in a self-paced manner (i.e., each sentence remained on the screen until the participant pressed a key) The interval between the disappearance of a sentence and the presentation of the next one was 500 msec. One line of each story was presented word-by-word in the center of the screen and each word appeared for 800 msec. In both the *Ironic* and *Literal* conditions the target line was always presented word-by-word. On the contrary, in the other 60 stories the line presented word-by-word was chosen randomly (see Figure 1) in order to prevent word-by-word presentation from being a strong cue that indicates which is the target line. After the last sentence (line 7) disappeared, the question was presented in one third of the trials and the participants pressed one of two buttons on a mouse (yes/no response) to answer (see Figure 1). Variable periods of visual fixation (1000 msec. + between 1 and 1000 msec.) were added at the end of each trial. The presentation order of the stories was pseudo-randomized. This means that the number of ironic and literal stories, decoys and fillers was balanced among the sessions. 15 ironic stories, 15 literal stories, 15 decoys and 15 positive fillers were presented in the first half of the experiment and 15 ironic stories, 15 literal stories, 15 decoys and 15 positive fillers were presented in the second half. Each stimulus was displayed in a left-justified manner. Participants were instructed to read at a normal rate and to respond as accurately as possible to the questions. The experimental session began with 2 training trials, which did not include ironies. All told, a typical session lasted an hour (including breaks).

Insert Figure 1 about here

2.4   Electroencephalogram (EEG) recording

EEG data were recorded using BrainAmp amplifiers (Brain Vision recorder software, Brain Products GmbH, Munich, Germany). EEG was recorded from 32 scalp sites using the international 10-20 system (American Electroencephalographic Society, 1994), with a forehead ground. Impedance was 10 kΩ or less at the start of the recording. All scalp sites were referenced to the left mastoid. Horizontal and vertical eye movements were monitored

using electro-oculograms (EOG) with a bipolar recording from electrodes placed around the left eye. The signal was sampled at 500 Hz.

2.5 Analysis

Both the ERP and TF analyses were conducted using ELAN-Pack software developed at INSERM U821 (Lyon, France) (Aguera, Jerbi, Caclin, & Bertrand, 2011). Trials contaminated by eye blinks or eye movements (threshold: ± 75 µV) were not included in the analyses nor were trials affected by drifts (range value 150 µV and latency: 500 msec.). Data from 3 subjects were excluded from the analyses because noise and eye movements contaminated more than half of the trials in the target condition.

2.5.1 ERP analysis

ERP analysis consisted of averaging the EEG segments in synchronization with the onset of the last word of the sentence presented word-by-word over 1250 msec. (250 msec. of pre-onset and 1000 msec. post-onset). The signals were band-pass filtered (0.16 – 30 Hz) and a baseline correction was computed from the 250 msec. to the 50 msec. preceding the onset of the target word and a notch filter was applied (50 Hz).

Nine representative electrodes of the 10 – 20 system were chosen to define different scalp regions (frontal: F3, Fz and F4; central: C3, Cz and C4 and parietal: P3, Pz and P4). We ran multiple ANOVAs using repeated measures including *Condition* (*Ironic* or *Literal*) and two levels of Electrode Site: *Laterality* (Left, Midline and Right) and *Anterior–Posterior* location (Frontal, Central, Parietal) as within-subject factors. Relevant post hoc comparisons were computed with Tukey HSD tests. We concentrated the statistical analysis on the time window 300 – 500 msec. in order to test the presence of the N400 effect and on the time window 500 – 800 msec. for the P600.

2.5.2 TF analysis

Task-induced modulations of power across time and frequency were obtained by standard time–frequency (TF) analysis using wavelets (Tallon-Baudry, Bertrand, Delpuech, & Permier, 1997) over 1250 msec. (250 msec. of pre-onset and 1000 msec. post-onset of the last word of

the target sentence). A baseline correction was computed from the 250 msec. to the 50 msec. preceding the onset of the target word. We conducted statistical analyses over three frequencies bands: theta (4 – 7 Hz), alpha (8 – 12 Hz) and gamma (> 30 Hz). In the gamma band the analysis of the signal was band-pass filtered in multiple successive 5 Hz wide frequency bands (e.g., [31 – 35], [36 – 40]). To test for significant increases or decreases in a frequency band, we used a paired-sample Wilcoxon signed rank test followed by a false discovery rate (FDR) correction across all time samples. The FDR approach yields a corrected threshold for significance (Genovese, Lazar, & Nichols, 2002) (significant level $p < .05$).

3. Results

3.2 ERP results

Figure 2 shows the time-course of the ERPs elicited by the *Ironic* and the *Literal* conditions at Cz. The repeated measures ANOVA on the time window 300 – 500 msec. showed a main effect of both *Laterality* [$F(2, 32) = 4.43$, $p < .05$] and *Anterior-Posterior* [$F(2, 32) = 3.91$, $p < .05$] as a significant interaction between the two variables [$F(4, 64) = 2.62$, $p < .05$]. However, the statistical analysis showed no significant difference between the *Ironic* and the *Literal* conditions [$F(1, 16) = .03$, $p > .8$] nor any interaction between the variable *Condition* (*Ironic* / *Literal*) and the other variables. Therefore, our results do not reveal a N400 effect. In the time window 500 – 800 msec. the repeated measures ANOVA revealed a main effect of *Laterality* [$F(2, 32) = 11.23$, $p < .001$], a marginally significant effect of the variable *Condition* [$F(1, 16) = 3.61$, $p = .075$] and a significant interaction between *Condition* and *Anterior-Posterior* [$F(2, 32) = 5.84$, $p < .01$]. Through a Tukey HSD test, one can see that the difference between the *Ironic* and the *Literal* conditions were due to a positive enhancement of the ERPs wave in the *Ironic* condition that were superior in the frontal sites. Considering the polarity and the shape of the ERP wave for the *Ironic* condition, we tend to consider this difference in enhancement as a P600 effect (see Figure 2).

Insert Figure 2 about here

3.3 TF results

Visual inspection of the spectrum of frequencies suggested the presence of an increase of power both in the theta band (4 – 7 Hz) and in the gamma band (30 – 90 Hz) during the *Ironic* condition. In addition, the power in the alpha band (8 – 12 Hz) seemed to increase in the frontal areas and to decrease in the parietal regions in this condition. The statistical analysis revealed that the difference of power in the theta band between the *Ironic* and the *Literal* conditions is significant between 500 and 700 msec., especially in the right frontal regions of the scalp. In the alpha band, comparisons between the *Ironic* and the *Literal* conditions were significantly different between (a) 400 and 700 msec. in the right frontal part of the scalp and between; (b) 550 and 700 msec. in the left parietal areas. However, spot (a) indicates an increase of power in the alpha band while spot (b) reveals a decrease of power in the same band. The statistical analysis also showed a significant increase of power in the early range (31 – 35 Hz) of the gamma band between 280 and 400 msec. in the frontal areas of the scalp (see Figure 3 and 4).

Insert Figure 3 and 4 about here

----

4. Discussion

Understanding a sentence is a complex job and pragmatic inferences are often necessary in order to grasp the complete meaning of an utterance. The present study was designed to investigate the allocation of different cognitive resources during pragmatic inference making with respect to irony. Irony is our test bed because it represents a clear case in which the linguistic code underdetermines the complete meaning of the utterance and so pragmatic inferences are necessary to fill in the gap between the code and the interpretation of the ironic remark. We tried to combine information both from ERPs and TF analysis to shed some light on pragmatic inference making.

4.1 ERP analysis

The ERP analysis focused on the N400 and the P600 components because the literature shows that these two components should be highly sensitive to pragmatic inference making. The classic decoded-oriented view of language comprehension would lead one to expect an N400 effect for the contrast *Ironic > Literal* because the literal meaning of an ironic remark is banally false. However, our results do not support this prediction, showing no significant increase of the N400 component due to irony. The absence of an N400 effect is in line with the results of the most recent ERP studies on irony processing (Regel et al., 2010, 2011). We do not want to go too far with the interpretation of a null result but the absence of a N400 effect suggests that the semantic/world knowledge violation is not perceived as a critical aspect of irony processing. Why not? Consider, for example, the study by Hagoort et al. (2004) in which (3b) ("The Dutch trains are *white* and very crowded.") is inconsistent only with the common experience of Dutch people, and a new line of Dutch trains could absolutely be white. Nevertheless, (3b) elicits an increase in the N400 component when compared with (3a). What is the difference between (3b) and the ironic interpretation of (1)? In (3b), the violation of world knowledge prevents the sentence from being integrated into the context (i.e., the subject's representation of the state of affairs), while when (1) is preceded by an appropriate context a surface violation only discloses the ironic interpretation as it triggers pragmatic inferences. Therefore, in (3b) the violation blocks the process of interpretation, but in (1) the surface contrast with the context allows one to go beyond the linguistic code (see Regel et al., 2011 for a similar interpretation).

The *ironic* condition elicited a strong positive enhancement of the ERP curves from 500 msec. and onward when compared to the *Literal* condition. The shape and latency of the ERP curves allows us to consider this enhancement a P600 effect. As previously mentioned, it is hard to provide a functional interpretation of the P600, but Regel and colleagues (2011) have proposed several possible interpretations. First, the P600 can be modulated by the predictability of the stimuli and by the experimental task (see for example, Fischler & Bradley, 2006; Kanske & Kotz, 2007). Second, the P600 might index the processing of the emotional arousal caused by the ironic stimuli because it has been found that arousal pictures and emotional words elicit an increase of the P600 in comparison with neutral controls (see e.g., Fischler & Bradley, 2006; Herbert, Junghofer, & Kissler, 2008; Kissler, Herbert, Winkler, & Junghofer, 2009). Third, P600 modulation may reflect the processing of pragmatic inferences and, critically, the "reintegration of semantic meaning with

extralinguistic information" (Lattner & Friederici, 2003). Regel et al. (2011) tend to reject the first and second interpretations due to the structure of their studies and we do not think that those hypotheses fit with the present study either. First of all, the presence of decoys prevents negative contexts from being reliable cues to ironic remarks. In addition, the task consisted of answering an easy comprehension question that never concerned the target line and it was present only in one third of the trials. Therefore, we believe that the enhancement of the P600 should not be due to the structure of the design. The level of arousal caused by the stimuli may modulate the P600 but we do not believe that this is the case in our study because the contrast *Ironic* vs *Literal* does not oppose an emotionally-charged remark and a neutral sentence. The target sentence, which is the same in both the *Ironic* and the *Literal* conditions, expresses an evaluation in both cases and so the contrast *Ironic* vs *Literal* opposes sentences with a comparable level of arousal. We are sympathetic with Regel et al.'s (2011) proposal that P600 may reflect the integration of the linguistic stimulus with extralinguistic information. In addition, several studies have shown that the P600 is modulated by the effort to integrate the target sentence into the discourse context (e.g., Nieuwland & Berkum, 2005).

In a review on the N400 and P600 components, Gina Kuperberg (2007) provides very interesting insights into the possible mechanisms that could generate a P600 effect. She proposed that language processing engages at least two routes to comprehend an utterance. The first is a semantic memory-based system that compares lexical information about the incoming words with information that is already stored in the semantic memory, and the work of that system should be reflected by the N400. The second route would be a combinatorial process that integrates words to build up the propositional meaning on the basis of multiple constraints. She proposed that the P600 might reflect the continued combinatorial analysis, which, in the end, determines the final interpretation of the sentence. Though her approach is centered on morphosyntactic and thematic–semantic constraints, the last line of the paper seems open to a broader interpretation of the mechanisms she proposed:

> The idea that there are multiple distinct but interactive processing streams underlying comprehension helps explain how, on the one hand, we make maximal use of what we have encountered again and again in the real world, and yet how, on the other hand, we are able to compute unusual relationships between people, objects and actions to understand novel events. The balanced operation of these distinct brain systems – one that links incoming semantic information with existing information stored in

semantic memory, and another that combines relationships between people, objects and actions to construct new meaning – allows for comprehension that is both efficient and yet adaptive. (Kuperberg, 2007, p. 45)

Irony processing clearly requires one to construct a new meaning that does not match the simple combination of the lexical meanings of the incoming words. As we have already underlined, the addressee has to go beyond the linguistic code by also engaging extralinguistic resources, namely Theory of Mind. Therefore, the increase of the P600 during the *Ironic* condition may index the continued combinatory analysis among different cognitive resources that leads to the interpretation of the ironic remark. In addition, studies that contrasted false belief and true belief stories have shown an enhancement of a positive late slow wave (Meinhardt et al., in press; Geangu et al., 2012; Meinhardt et al., 2011; Sabbagh & Taylor, 2000). However, we would like to go further in the investigation of pragmatic inference making and, in order to accomplish this aim, we conducted the time frequency analysis.

4.2 TF analysis

As we have already remarked, the literature on brain oscillations during language processing does not yet provide a clear rubric with which one can easily interpret new data, and so the interpretation of our TF results is exploratory. Like Klimesch and colleagues (e.g., 1996; 1999), we claim that the increase of power in the theta band reflects the encoding of new information into episodic memory. This result is particularly interesting because the target sentence is exactly the same in both the *Ironic* and *Literal* conditions and so the extra meaning, which has to be encoded, is attributed to pragmatic inference. Consider once more the scenario of the two opera singers. If they gave a great performance, (4) communicates that the speaker is happy, while if they gave a bad performance, (4) communicates that the speaker is partially angry, partially sad (and other negative attitudes), and additionally allows the speaker to be a bit cutting or even accusatory in her remarks (for the range of extra effects an ironic statement can cause see for example Haverkate, 1990; Katz & Lee, 1993). In the end, the payoff of using irony instead of a direct remark is that irony is more informative because it conveys a broader range of cognitive effects than a literal remark. The encoding of this extra information strongly engages short-term episodic memory, which is detected by the increase of power in theta band.

Alpha waves have been known to reflect attentive processes since the pioneering work by Dement and Kleitman (1957) in which they showed that the onset of alpha power decreases during the transition from alert wakefulness to sleep. However, more recent studies (e.g., Klimesch, Doppelmayr, Pachinger, et al., 1997; Klimesch, Doppelmayr, Schimke, et al., 1997; Röhm, Klimesch, Haider, & Doppelmayr, 2001) have revealed that a decrease of power in the alpha band, especially in the upper range (10 – 12 Hz), is related to semantic (long term) memory. Both hypotheses make sense in the context of irony processing but we believe that the link between desynchronization in the alpha band and attentive processes is more relevant. The computation of pragmatic inferences is demanding and so attentive resources are recruited as the cognitive load increases. Our data also revealed an increase of power in the alpha band between 400 and 700 msec. and the synchronization of the alpha waves is considered to indicate a state of 'idling' (e.g., Pfurtscheller, Stancak, & Neuper, 1996), or even a state of inhibition (Klimesch, 1996) of cortical regions that are not involved in the current processes. We can say that, during a demanding task, cognitive resources are allocated to the systems that require them and Klimesch, Doppelmayr, Schwaiger, Auinger, & Winkler (1999) argued, in particular, that synchronization of alpha waves correlates with processes that maximize episodic short-term memory. We have claimed that the synchronization of theta waves indexes the encoding of new information in episodic memory and the increase of power in the alpha band during roughly the same time window reinforces our hypothesis.

All the information made available by the incoming words and by the context has to be integrated into a meaningful interpretation of the utterance and researchers such as Hagoort and Bastiaansen argue that synchronization in the gamma band reflects these "unification operations" in the semantic domain. However, as Bastiaansen et al. (2012) underlined, research is still limited to the investigation of the unification processes at the syntactic and semantic levels. Our study broadens the research spectrum by investigating a prototypical case of pragmatic inference making and the data still show an increase of power in the low gamma band during the *Ironic* condition when compared to the *Literal* one. Our results reveal that gamma waves reflect unification operations in both semantic and pragmatic domains, but we hope to have gone a bit further.

As we said in the beginning, pragmatic inference making critically requires one to access the speaker's intentions and this process involves cognitive resources that likely pertain to the domain of social cognition, namely ToM. There is already quite a bit of evidence that the gamma band can index social dimensions of communication (e.g., Grossmann et al., 2008; van den Brink et al., 2010). Spotorno et al. (2012) has shown that the

interaction between the linguistic and ToM networks plays a crucial role in irony processing and we argue here that the interaction between those kinds of cognitive abilities is reflected by an increase of power in the gamma band during pragmatic inference making.

5. Conclusion

The present study was designed to investigate the cognitive processes behind pragmatic inference making. We chose irony as our test bed because irony processing obviously requires one to go beyond the linguistic code to grasp the complete meaning of the ironic remark and pragmatic inferences are critical to filling in the gap between the code and the complete meaning of the utterance. The ERP results largely replicate the main effects of the most recent studies on irony (e.g., Regel et al., 2010; 2011); namely, an increase of the P600 during the *Ironic* condition when compared to the *Literal* one. In line with Regel et al.'s (2011) interpretation, we argue that the P600 reflects effort in the integration processes between the ironic utterance and the preceding context. However, in order to dig deeper into the investigation of pragmatic inference making, we also conducted a time frequency analysis. The picture that emerges from the TF analysis is that irony processing is highly demanding in terms of attentional resources, as shown by the decrease of power in the alpha band, and that it also extensively recruits short-term memory in order to encode new information into episodic memory (as supported by the increase of power in roughly the same time window of both the theta band and the alpha band). In addition, irony processing requires cognitive systems to coordinate and unify several streams of information. We argue that the increase of power in the gamma band reflects the integration between different cognitive resources, namely linguistic and ToM abilities.

Our reading of the TF data is generally in line with the interpretation provided by Bastiaansen and colleagues (2012), but we tried to extend the analysis beyond semantics by integrating the pragmatic dimension. The present study is based on the investigation of a specific linguistic phenomenon, namely irony, but irony processing is just a prototypical case of an instance in which one has to go beyond the linguistic code to grasp the complete meaning of the utterance. Therefore, we propose that the same kinds of variations in the theta, alpha and gamma bands, and the tentative interpretation that we provided for them may be applied to pragmatic inference making in general.

6. References

Aguera, P.-E., Jerbi, K., Caclin, A., & Bertrand, O. (2011). ELAN: a software package for analysis and visualization of MEG, EEG, and LFP signals. *Computational intelligence and neuroscience*, 1–11.

Amenta, S., & Balconi, M. (2008). Understanding irony: An ERP analysis on the elaboration of acoustic ironic statements. *Neuropsychological Trends*, *3*, 7–27.

American Electroencephalographic Society, . (1994). American Electroencephalo- graphic Society. Guideline thirteen: Guidelines for standard electrode position nomenclature. *Journal of Clinical Neurophysiology*, *11*, 111–113.

Bastiaansen, M. C. M., Oostenveld, R., Jensen, O., & Hagoort, P. (2008). I see what you mean: theta power increases are involved in the retrieval of lexical semantic information. *Brain and language*, *106*(1), 15–28.

Bastiaansen, M. C. M., van der Linden, M., Ter Keurs, M., Dijkstra, T., & Hagoort, P. (2005). Theta responses are involved in lexical-semantic retrieval during language processing. *Journal of cognitive neuroscience*, *17*(3), 530–41.

Bastiaansen, M., Mazaheri, A., & Jensen, O. (2012). Beyond ERP's: oscillatory neuronal dynamics. In S. J. Luck & E. S. Kappenman (Eds.), *The Oxford handbook of event-related potential components* (pp. 31–50). New York, NY: Oxford University Press.

Bloom, P. (2000). *How Children Learn the Meanings of Words*. Cambridge: MIT Press.

Bornkessel-schlesewsky, I., & Schlesewsky, M. (2008). An alternative perspective on " semantic P600 " effects in language comprehension. *Brain Research Reviews*, *59*(1), 55–73.

Clark, H. H., & Gerrig, R. J. (1984). On the pretense theory of irony. *Journal of experimental psychology. General*, *113*(1), 121–6.

Cornejol, C., Simonetti, F., Aldunate, N., Ibáñez, A., López, V., & Melloni, L. (2007). Electrophysiological evidence of different interpretative strategies in irony comprehension. *Journal of psycholinguistic research*, *36*(6), 411–30.

Coulson, S., & Lovett, C. (2010). Comprehension of non-conventional indirect requests: An event-related brain potential study. *Italian Journal of Linguistics*, *22*(1), 107–124.

Dement, W., & Kleitman, N. (1957). Cyclic variations in EEG during sleep and their relation to eye movements, body motility, and dreaming. *Electroencephalography and Clinical Neurophysiology*, *9*(4), 673–690.

Ferstl, E. C., Neumann, J., Bogler, C., & Von Cramon, D. Y. (2008). The extended language network: A meta-analysis of neuroimaging studies on text comprehension. *Human brain mapping*, *29*(5), 581.

Filippova, E., & Astington, J. W. (2008). Further Development in Social Reasoning Revealed in Discourse Irony Understanding. *Memory*, *79*(1), 126 – 138.

Fischler, I., & Bradley, M. (2006). Event-related potential studies of language and emotion: words, phrases, and task effects. *Progress in brain research: Understanding Emotions*, *156*.

Friederici, A., & Meyer, M. (2004). The brain knows the difference: Two types of grammatical violations. *Brain research*, *1000*(1-2), 72–77.

Genovese, C. R., Lazar, N. A., & Nichols, T. (2002). Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *NeuroImage*, *15*(4), 870–878.

Grauwe, S. D., Swain, A., & Holcomb, P. (2010). Electrophysiological insights into the processing of nominal metaphors. *Neuropsychologia*, *48*(7), 1965–1984.

Grice, P. (1975). Logic and Conversation. In P. Cole & Morgan J. (Eds.), *Syntax and Semantics* (pp. 41 – 58). Ney York: Academic Press.

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

Grossmann, T., Johnson, M. H., Lloyd-Fox, S., Blasi, A., Deligianni, F., Elwell, C., & Csibra, G. (2008). Early cortical specialization for face-to-face communication in human infants. *Proceedings of The Royal Society*, *275*(1653), 2803–11.

Hagoort, P. (2005). On Broca, brain, and binding: a new framework. *Trends in cognitive sciences*, *9*(9), 416–23.

Hagoort, P., Brown, C. M., & Osterhout, L. (1999). The neurocognition of syntactic processing. In C. Brown & P. Hagoort (Eds.), *Neurocognition of Language* (pp. 273–316). Oxford: Oxford University Press.

Hagoort, P., Hald, L., Bastiaansen, M., & Petersson, K. M. (2004). Integration of word meaning and world knowledge in language comprehension. *Science*, *304*(5669), 438–41.

Happé, F. G. (1993). Communicative competence and theory of mind in autism: a test of relevance theory. *Cognition*, *48*(2), 101–119.

Haverkate, H. (1990). A speech act analysis of irony. *Journal of Pragmatics*, *14*(1), 77–109.

Herbert, C., Junghofer, M., & Kissler, J. (2008). Event related potentials to emotional adjectives during reading. *Psychophysiology*, *45*(3), 487–98.

Herten, M. V., Kolk, H. H. J., & Chwilla, D. J. (2005). An ERP study of P600 effects elicited by semantic anomalies. *Cognitive Brain Research*, *22*, 241 – 255.

Kanske, P., & Kotz, S. a. (2007). Concreteness in emotional words: ERP evidence from a hemifield study. *Brain research*, *1148*, 138–48.

Katz, A., Blasko, D., & Kazmerski, V. (2004). Saying what you don't mean social influences on sarcastic language processing. *Current Directions in Psycological Science*, *13*(5), 186–189.

Katz, A., & Lee, C. (1993). The Role of Authorial Intent in Determining Verbal Irony and Metaphor. *Metaphor and Symbol*, *8*(4), 257–279.

Kissler, J., Herbert, C., Winkler, I., & Junghofer, M. (2009). Emotion and attention in visual word processing: an ERP study. *Biological Psychology*, *80*(1), 75–83.

Klimesch, W. (1994). *The structure of long-term memory: A connectivity model of semantic processing*. Hillsdale, NJ: Lawrence Erlbaum.

Klimesch, W. (1996). Memory processes, brain oscillations and EEG synchronization. *International Journal of Psychophysiology*, *24*(1-2), 61–100.

Klimesch, W. (1999). EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. *Brain research reviews*, *29*(2-3), 169–195.

Klimesch, W., Doppelmayr, M., Pachinger, T., & Ripper, B. (1997). Brain oscillations and human memory: EEG correlates in the upper alpha and theta band. *Neuroscience letters*, *238*(1-2), 9–12.

Klimesch, W., Doppelmayr, M., Schimke, H., & Ripper, B. (1997). Theta synchronization and alpha desynchronization in a memory task. *Psychophysiology*, *34*(2), 169–176.

Klimesch, W., Doppelmayr, M., Schwaiger, J., Auinger, P., & Winkler, T. (1999). "Paradoxical" alpha synchronization in a memory task. *Cognitive brain research*, *7*(4), 493–501.

Kuperberg, G. (2007). Neural mechanisms of language comprehension: Challenges to syntax. *Brain Research*, *1146*, 23–49.

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, *207*(4427), 203–205.

Langdon, R., Davies, M., & Coltheart, M. (2002). Understanding minds and understanding communicated meanings in schizophrenia. *Mind & Language*, *17*(1-2), 68–104.

Lattner, S., & Friederici, A. D. (2003). Talker's voice and gender stereotype in human auditory sentence processing – evidence from event-related brain potentials. *Neuroscience Letters*, *339*(3), 191–194.

Nieuwland, M. S., & Berkum, J. J. A. V. (2005). Testing the limits of the semantic illusion phenomenon : ERPs reveal temporary semantic change deafness in discourse comprehension. *Cognitive Brain Research*, *24*(3), 691 – 701.

Nieuwland, M. S., 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*(3), 324–346.

Nieuwland, M. S., & Van Berkum, J. (2006). When peanuts fall in love: N400 evidence for the power of discourse. *Journal of Cognitive Neuroscience*, *18*(7), 1098–1111.

Noveck, I. a, & Posada, A. (2003). Characterizing the time course of an implicature: An evoked potentials study. *Brain and Language*, *85*(2), 203–210.

Pfurtscheller, G., & Lopes da Silva, F. H. E. (1999). Event-related desynchronization. *Handbook of Electroencephalography and Clinical Neurophysiology, revised series (Vol. 6)*. Amsterdam: Elsevier.

Pfurtscheller, G., Stancak, A., & Neuper, C. (1996). Event-related synchronization (ERS) in the alpha band—an electrophysiological correlate of cortical idling: a review. *International Journal of Psychophysiology*, *24*(1-2), 39–46.

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a Theory of Mind? *Behavioral and brain sciences*, *1*(4), 515–526.

Regel, S., Coulson, S., & Gunter, T. C. (2010). The communicative style of a speaker can affect language comprehension? ERP evidence from the comprehension of irony. *Brain research*, *1311*, 121–35.

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*(2), 277–293.

Röhm, D., Klimesch, W., Haider, H., & Doppelmayr, M. (2001). The role of theta and alpha oscillations for language comprehension in the human electroencephalogram. *Neuroscience Letters*, *310*(2-3), 137–140.

Singer, W. (1993). Synchronization of cortical activity. *Annual Rev of Psychophysiologyv of Psychophysiology*, *55*, 349–374.

Singer, Wolf. (1999). Neuronal Synchrony : A Versatile Code for the Definition of Relations ? Most of our knowledge about the functional organization. *Neuron*, *24*(1), 49–65.

Sperber, D., & Wilson, D. (1986). *Relevance: Communication and Cognition*. Oxford: Blackwell; Wiley-Blackwell; 2nd Edition edition 1995.

Spotorno, N., Koun, E., Prado, J., Van Der Henst, J.-B., & Noveck, I. A. (2012). Neural evidence that utterance-processing entails mentalizing: The case of irony. *NeuroImage*, *63*(1), 25–39.

Sterman, M., Kaiser, D., & Veigel, B. (1996). Spectral Analysis of Event-Related EEG Responses During Short-Term Memory Performance. *Brain topography*, *9*(1), 21–30.

Tallon-Baudry, C., Bertrand, O., Delpuech, C., & Permier, J. (1997). Oscillatory gamma-band (30-70 Hz) activity induced by a visual search task in humans. *The Journal of neuroscience*, *17*(2), 722–734.

Van Overwalle, F., & Baetens, K. (2009). Understanding others' actions and goals by mirror and mentalizing systems: a meta-analysis. *NeuroImage*, *48*(3), 564–84.

Varela, F., & Lachaux, J. (2001). The brainweb: phase synchronization and large-scale integration. *Nature reviews Neuroscience*, *2*(April).

Weiss, S., Rappelsberger, P., Schack, B., & Mueller, H. M. (2003). Koharenz- und Phasenuntersuchungen und ihre Bedeutung fur die Untersuchung von Sprachprozessen. In G. R. H. M. Mueller (Ed.), *Neurokognition der Sprache*. Tubingen: Stauffenburg.

Weiss, Sabine, Mueller, H. M., Schack, B., King, J. W., Kutas, M., & Rappelsberger, P. (2005). Increased neuronal communication accompanying sentence comprehension. *International journal of psychophysiology*, *57*(2), 129–41.

Wilson, D., & Sperber, D. (2012). Explaining Irony. In D. Wilson & D. Sperber (Eds.), *Meaning and Relevance*. Cambridge: Cambridge University Press.

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, *13*(1), 103–128.

van Berkum, J. J., Zwitserlood, P., Bastiaansen, M., Brown, C., & Hagoort, P. (2004). So who's "he" anyway? Differential ERP and ERSP effects of referential success, ambiguity and failure during spoken language comprehension. *Paper presented at the Annual Meeting of the Cognitive Neuroscience Society, San Francisco*.

van den Brink, D., Van Berkum, J. J. a, Bastiaansen, M. C. M., Tesink, C. M. J. Y., Kos, M., Buitelaar, J. K., & Hagoort, P. (2010). Empathy matters: ERP evidence for inter-individual differences in social language processing. *Social cognitive and affective neuroscience*, *7*(2), 173–83.

Table 1.

An example from the stimuli (translated from French).

Condition	Example
Ironic	Cynthia and Léa sing together in the same opera. On the night of the premiere they meet at the theatre. The show begins exactly on time. During their performance both ladies sing off key. After the show, Cynthia says to Léa: "Tonight we gave a superb performance." As they take off their make-up they continue to discuss the show. Question: Do you think that the performance was in the morning?
Literal	Cynthia and Léa sing together in the same opera. On the night of the premiere they meet at the theatre. The show begins exactly on time. Both ladies sing beautifully and receive a rapturous round of applause. After the show, Cynthia says to Léa: "Tonight we gave a superb performance" As they take off their make-up they continue to discuss the show. Question: Do you think that the performance was in the morning?
Decoy	Mateo is relocating and has to move a very fragile and heavy mirror. He asks Paul for help. Paul makes himself available immediately. As soon as Paul lifts the mirror it breaks into a thousand pieces. Mateo says to Paul: "We have made a big mistake." A few days later, Mateo celebrates his move with his friends. Question: In your opinion, do Mateo and Damien move the mirror without problems?
Filler	Jeremy has promised to his kid to build him a cabin. He bought chestnut wood to build it. He works all the afternoon to finish it. In the end, the cabin is solid and well built. His kid is very happy and he tells him: "Come to play with my in the cabin." They play all the weekend long in this new cabin. Question: Do you think that the cabin is well built?

Figure legends

Figure 1. Experimental procedure.

Figure 2. ERP waves for the *Ironic* (red line) and *Literal* (blue line) conditions at Cz. The gray rectangle highlights the significant interval. On the right there are the scalp distribution for both conditions at a significant latency for the P600 component.

Figure 3. From the left: Z value for the contrast *Ironic < Literal* in the theta band at F4; Z value for the contrast *Ironic < Literal* in the alpha band at F4; Z value for the contrast *Ironic < Literal* in the alpha band at P3 and Z value for the contrast *Ironic < Literal* in the gamma band at Fz.

Figure 4. From the left: scalp distribution of theta waves (Z value for the contrast *Ironic < Literal*); scalp distribution of alpha waves (Z value for the contrast *Ironic < Literal*) and scalp distribution of gamma waves (Z value for the contrast *Ironic < Literal*).
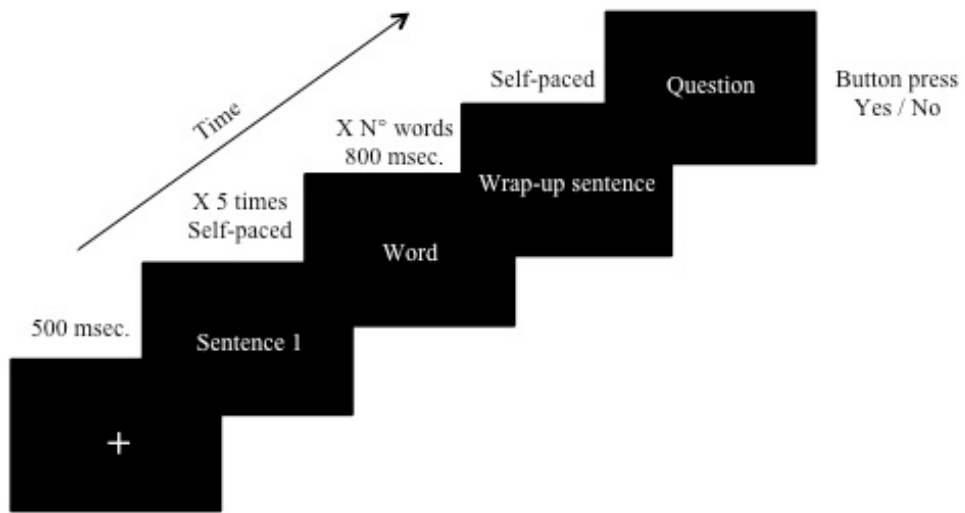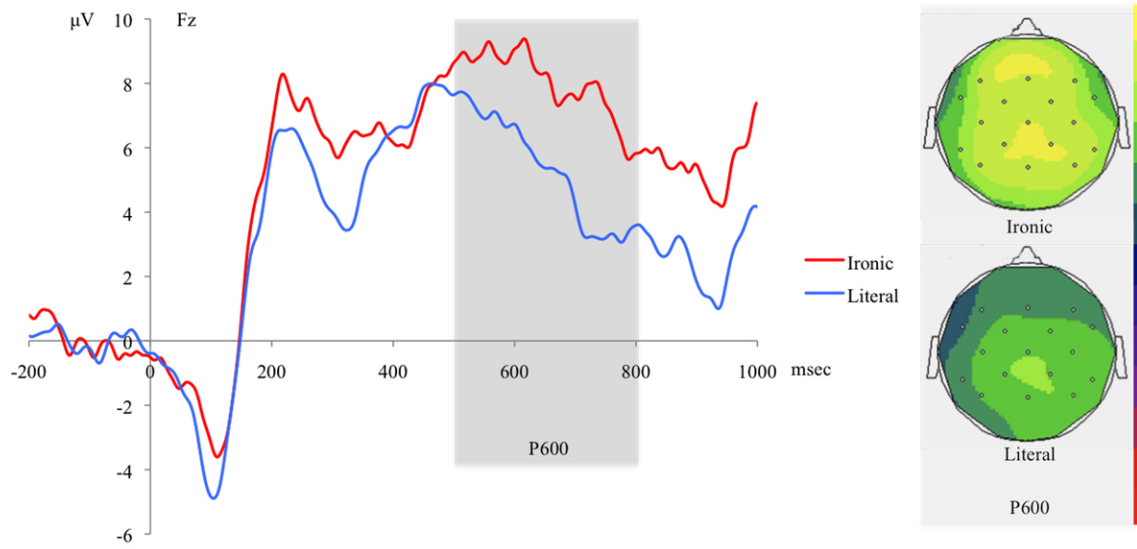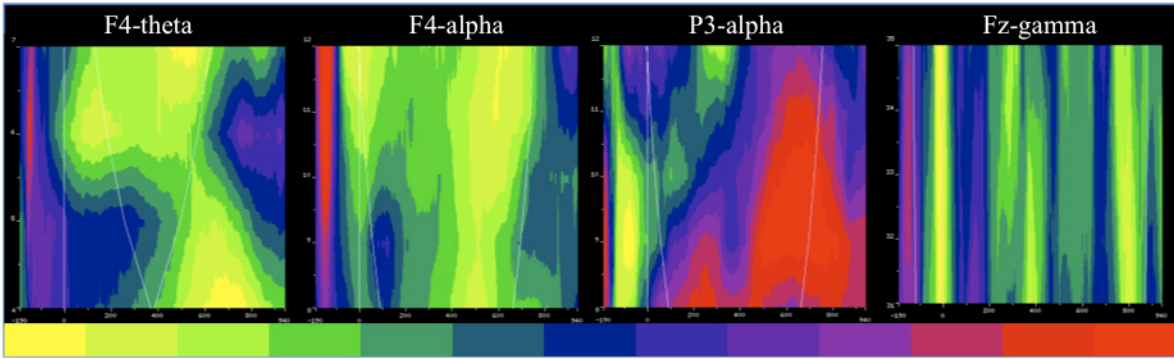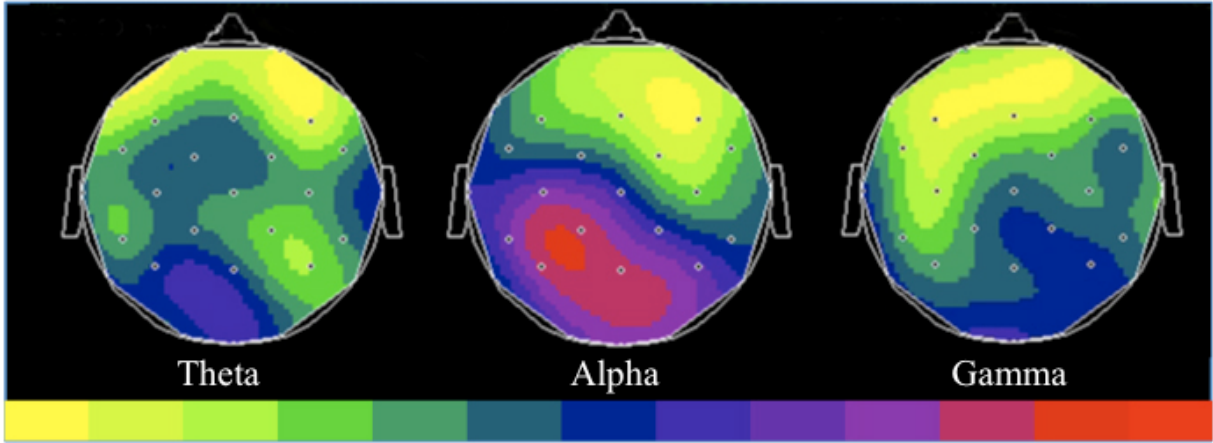
Figure 1.

Figure 2.

Figure 3.

Figure 4.

CONCLUDING REMARKS

The present work was designed to investigate the cognitive and neural bases of irony processing because we consider irony a perfect object of study for Experimental Pragmatics. Irony shows, first, how a given utterance can have a thoroughly different meaning as a function of context. Second, it reveals that processes other than the encoding-decoding machinery are involved in language comprehension. Third, the target sentence can itself be the control. As a result, it provides a perfect test bed to investigate the interaction between language and Theory of Mind.

However, why should irony processing engage "more" ToM work than other forms of figurative language, e.g. metaphor? It is generally accepted that metaphor processing engages first level ToM while irony processing should require second order ToM (Happé, 1993). The so-called first order ToM is the ability to interpret others' thoughts and it is engaged by the classic false beliefs tasks, while second order ToM is the ability to interpret the thoughts of someone that is thinking about the thoughts of someone else. Second order ToM should be engaged in the resolution of more complex versions of the FBT in which, for example, Anne changes the location of the marble thinking that Sally cannot see, but Sally, looking from the peephole of the door, is observing the action. When Sally returns to the scene, the subject is asked where Anne thinks that Sally will look for the marble (e.g., Happé, 1993; Sperber and Wilson 1981) and so the subject has to be able to reflect on the beliefs of a character about the beliefs of another character. Therefore, irony provides a real advantage in the study of the interaction between language and Theory of Mind because it maximizes the involvement of ToM ability in language comprehension.

9.1 The cognitive and neural basis of irony processing

Irony processing has already been the object of empirical investigation but the research has rarely been focused on attitude ascription. The work here represents an advance in the following three ways. First of all, we brought attitude ascription back to center stage with a series of reading time experiments. Our data have shown that the comprehension of irony is effortful, but that the cognitive costs of irony processing can be dramatically reduced over time and if a cue, like the to-be-echoed prime, leads the addressee to go beyond the linguistic code in order to grasp the speaker's attitude. Behavioral studies also allowed us to set up important methodological constraints. The manipulation of the proportion of target items to fillers contributes to speeding up or slowing down irony processing and the choice of

the right control conditions affects the results of the experiments (see experiments 1 and 2). We introduced decoy stories (stories in which a negative context leads to a banal, non-ironic sentence) in order to prevent negative contexts from being cues of the presence of ironies and in Experiment 2 revealed that the presence of decoys blocks the Early-Late effect we found in the first experiment.

Second, the investigation of the neural bases of irony processing and the imaging studies have taken advantage of the methodological constraints of the reading time experiments. In the fMRI study, both the whole brain analysis and the analysis based on regions of interest (ROIs) have revealed that the entire bilateral ToM network is more active in the Ironic condition when compared with the Literal one. The ROIs – in the rTPJ, lTPJ, MPFC and PC — were chosen from a recent and extensive meta-analysis of the neuroimaging literature on ToM. Therefore, we can say that we investigated the engagement of the ToM network during irony processing by testing the activation of the most prototypical areas for ToM processing. In addition, the analyses of functional connectivity, through Psychophysiological interaction (PPI) analysis, suggested that information shared by the MPFC and with the left IFG is crucial for linking ToM activity with language processing. These results are in stark contrast with previous studies on verbal irony, which have found very little pointing to ToM activation.

At this point, our investigation of irony processing was still lacking information about the allocation of the different cognitive resources during the on-line processing of an irony and we tried to fill in the gap with the EEG study. The third advance can be seen through ERP analyses which showed an enhancement of the P600 component for the contrast *Ironic > Literal*. This effect, which has already been observed in previous studies on irony (e.g., Regel et al. 2010; 2011), can be associated both with integration operations in the domain of discourse processing (e.g., Nieuwland & van Berkum, 2005) and with the integration between linguistic and extralinguistic information (Regel et al., 2011). Especially novel was the Time Frequency analysis which revealed that the comprehension of an ironic remark is demanding in terms short-term memory, due to the encoding of new information into the episodic memory. These data are particularly interesting considering that in both the Ironic and the Literal conditions the target sentence is exactly the same, and so the extra information, which requires more work in the short-term memory, derives from pragmatic inferences. The TF analysis also showed an increase of power in the gamma band that probably sustains the integration operations among different streams of information and different cognitive resources from the first steps of the comprehension process.

The behavioral, fMRI and EEG studies have provided us with a rough sketch of the cognitive and neural bases of irony processing. Each study looked the same phenomenon employing the same basic design, but from a different angle. The integration of several techniques allowed us to partially overcome the limits of each of them.

9.2 <u>Contributions to the debate on irony processing</u>

What makes irony understanding different from the understanding of a literal remark? This is arguably the motivating question behind every study on irony processing in the experimental literature. In this respect, the most evident contribution of the thesis has been to show that attitude ascription is an essential factor and that to ignore ToM is to ignore the hallmark of irony. The thesis has, also, argued that accounts such as the Direct Access View, the Graded Salience Hypothesis or those based on the SPM, ignore attitude ascription by focusing their attention on the alleged priority of a literal interpretation in an ironic remark; nevertheless, those accounts propose interesting predictions on irony processing and any researcher needs to face up to them.

The thesis clearly supports the hypothesis that irony understanding is complex and effortful: it engages an extensive brain network and it is demanding in terms of memory and integration operations. However, the behavioral experiments have revealed that irony processing is effortful only under certain specifiable conditions and that extra cues, such as those that use a to-be-echoed-prime, can make the comprehension process easier. Where does that leave the leading accounts? One can consider the current data as support for the Direct Access View because, above all, the main assumption behind Gibbs' account is that comprehending literal as well as non-literal meanings of a sentence largely depends on pragmatic knowledge and listeners' figurative modes of thought (Gibbs, 1994, 2002). Nevertheless, the context-dependent-nature of irony calls for a less generalized interpretation; the results of the second behavioral study (i.e., longer reading times for ironies with respect to literal counterparts) are not experimental artifacts: it is entirely possible that in everyday life an effective ironic remark comes up without strong background cues. In that case, irony processing would arguably require more effort than the process of a literal sentence. In addition, the ERPs results showed that irony elicits a significant increase of the P600. Considering that the enhancement of an ERP component is arguably the most sensitive index of a reaction of one's cognitive systems to an experimental manipulation (see for example

Van Berkum, 2004), the P600 effect argues against an equivalence between the comprehension of an ironic and a literal remark. On the other hand, predictions based on the SPM seem to be knocked out by the methodological issues that are at the basis of this proposal. As I have already argued, the direct translation of the Gricean theory into a rigid three-step psychological account for language processing has made it easy for critics to rail against the SPM, because such a three-step process seems too long and slow for explaining the rapid on-line pragmatic processing of an utterance.

The Graded Salience Hypothesis seems to be more flexible than the other two proposals, and, not by chance. Giora's account is often considered a reasonable explanation of both behavioral and neuro-imaging data (e.g., Filik & Moxey, 2010; Rapp et al., 2012; Regel et al. 2011). However, we claim that the role of salience in irony process is overestimated and the results presented in chapter 4 (the behavioral paper) support our hypothesis. The pool of stories was the same along the three experiments but the reading times have been dramatically affected by our subtle experimental manipulations which included the presence versus absence of decoys stories and the presence versus absence of the to-be-echoed prime and not the salience of the target sentences. Above all, our data have shown that ToM-revealing contextual aspects strongly affect irony processing while keeping constant the linguistic stimuli. These results are in contrast with the Graded Salience Hypothesis which predicts just a marginal effect of context. One of the take home messages of the thesis is that the debate on irony processing should return to the theories of Pragmatics about irony. By incorporating theoretical assumptions from pragmatics, one will be able to develop a more reliable model of irony comprehension and language processing in general.

9.3   To what extent can one generalize the present results?

One of the main claims in Pragmatics is that the linguistic code underdetermines the meaning of almost every utterance. In order to grasp the speaker's meaning one has to go beyond the linguistic code and, critically, she has to have access to the speaker's intentions. Theory of Mind is often necessary to fill in the gap between the linguistic code and the complete meaning, and so the interaction between language and Theory of Mind, which we have shown in the present work, should be common to every case of pragmatic inference making. Nevertheless, this hypothesis remains to be tested. Therefore, one of the main lines of development of the present work should be the investigation of other linguistic phenomena in

order to replicate our results on a broad range of cases. We started our research by employing a pragmatic phenomenon in which the communicative meaning is basically completely beyond the code, but by taking advantage of the insights that the studies presented here have provided the next challenge is to investigate to what extent our inferences can be generalized to every case of pragmatic inference making. Irony has served as a perfect test bed. While we think of it is just the tip of the iceberg, it remains to be seen just how deep it goes.

9.4   Other possible further directions

The studies that we have presented here provide a preliminary sketch of the interactions between language and Theory of Mind during the processing of an utterance, but this is undoubtedly just the beginning. Further investigation should a) deeply investigate the connections between language and ToM network and; b) verify to what extent our results can be replicated using different sorts of tests. One of the limits of the fMRI technique is that it cannot allow one to derive causal inferences from its results. That is, we cannot claim that the ToM process plays a causal role during irony processing. One way to bypass this obstacle would be to artificially knock out the ToM network and test whether irony processing is affected by this manipulation. Therefore, it would be eminently useful to conduct a study in which Transcranial Magnetic Stimulation (TMS) was applied to the critical ToM region. TMS is a non-invasive technique that artificially and temporarily disrupts the activity in a neural network. Therefore, TMS would allow one to test, for example, if the reading times for ironic and literal statements are differently affected by temporary blockage of the right TPJ.

A further direction of research would be to investigate the connectivity between the networks that are involved in pragmatic inference making. The PPI analysis reported in the fMRI study revealed an increase of functional connectivity between the mPFC and the left IFG during irony processing. This result can be considered a promising first step that suggests one ought to carry on study of both the functional and structural connectivity between the linguistic and ToM networks.

9.5   On ordering a beverage

In the Introduction to the thesis, I described a situation where I answered a linguistically-encoded request by relying exclusively on ToM. This showed how crucial ToM

is to communication generally. Much like with irony, successful communication depends on the listener making astute observations. Hopefully, future researchers who "hear" questions similar to those described in the thesis can also profit from the observable outcomes in the studies here.

REFERENCES

Aguera, P.-E., Jerbi, K., Caclin, A., & Bertrand, O. (2011). ELAN: a software package for analysis and visualization of MEG, EEG, and LFP signals. *Computational intelligence and neuroscience*, 1–11.

Amenta, S., & Balconi, M. (2008). Understanding irony: An ERP analysis on the elaboration of acoustic ironic statements. *Neuropsychological Trends*, *3*, 7–27.

American Electroencephalographic Society, . (1994). American Electroencephalo- graphic Society. Guideline thirteen: Guidelines for standard electrode position nomenclature. *Journal of Clinical Neurophysiology*, *11*, 111–113.

Apperly, I. a, Samson, D., Carroll, N., Hussain, S., & Humphreys, G. (2006). Intact first- and second-order false belief reasoning in a patient with severely impaired grammar. *Social neuroscience*, *1*(3-4), 334–438.

Astington, J. W., Pelletier, J., & Homer, B. (2002). Theory of mind and epistemological development: the relation between children's second-order false-belief understanding and their ability to reason about evidence. *New Ideas in Psychology*, *20*(2-3), 131–144.

Austin, J. L. (1962). *How to do Things with Words*. (J. O. Urmson, Ed.).

Baillargeon, R., Scott, R., & He, Z. (2010). False-belief understanding in infants. *Trends in cognitive sciences*, *14*(3), 110–118.

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*(3-4), 203–216.

Banerjee, R. (2000). The development of an understanding of modesty. *British Journal of Developmental Psychology*, *18*(4), 499–517.

Baron-Cohen, S. (1995). *Mindblindness: An essay on autism and theory of mind*. Boston: MIT Press/Bradford Books.

Baron-Cohen, S., & Wheelwright, S. (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*(1), 5–17.

Bastiaansen, M. C. M., Oostenveld, R., Jensen, O., & Hagoort, P. (2008). I see what you mean: theta power increases are involved in the retrieval of lexical semantic information. *Brain and language*, *106*(1), 15–28.

Bastiaansen, M. C. M., van der Linden, M., Ter Keurs, M., Dijkstra, T., & Hagoort, P. (2005). Theta responses are involved in lexical-semantic retrieval during language processing. *Journal of cognitive neuroscience*, *17*(3), 530–41.

Bastiaansen, M., Mazaheri, A., & Jensen, O. (2012). Beyond ERP's: oscillatory neuronal dynamics. In S. J. Luck & E. S. Kappenman (Eds.), *The Oxford handbook of event-related potential components* (pp. 31–50). New York, NY: Oxford University Press.

Bledowski, C., Prvulovic, D., Goebel, R., Zanella, F. E., & Linden, D. E. J. (2004). Attentional systems in target and distractor processing: a combined ERP and fMRI study. *NeuroImage*, *22*(2), 530–540.

Bloom, P. (2000). *How Children Learn the Meanings of Words*. Cambridge: MIT Press.

Bohrn, I. C., Altmann, U., & 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*(11), 2669–2683.

Bornkessel-schlesewsky, I., & Schlesewsky, M. (2008). An alternative perspective on " semantic P600 " effects in language comprehension. *Brain Research Reviews*, *59*(1), 55–73.

Bott, L., Bailey, T. M., & Grodner, D. (2012). Possibly all of that and then some: Scalar implicatures are understood in two steps. *Journal of Memory and Language*, *66*(1), 123–142.

Bott, L., & Noveck, I. (2004). Some utterances are underinformative: The onset and time course of scalar inferences. *Journal of memory and language*, *51*(3), 437–457.

Boulenger, V., Hauk, O., & Pulvermüller, F. (2009). Grasping ideas with the motor system: semantic somatotopy in idiom comprehension. *Cerebral cortex*, *19*(8), 1905–1914.

Brownell, H. H., Carroll, J. J., Rehak, A., & Wingfield, A. (1992). The use of pronoun anaphora and speaker mood in the interpretation of conversational utterances by right hemisphere brain-damaged patients. *Brain and language*, *43*(1), 121–47.

Brownell, H., Simpson, T., Bihrle, A., Potter, H. H., & Gardner, H. (1990). Appreciation of metaphoric alternative word meanings by left and right brain-damaged patients. *Neuropsychologia*, *28*(4), 375–383.

Burgess, C., & Chiarello, C. (1996). Neurocognitive mechanisms underlying metaphor comprehension and other figurative language. *Metaphor and symbol*, *11*(1), 67–84.

Buttelmann, D., Carpenter, M., & Tomasello, M. (2009). Eighteen-month-old infants show false belief understanding in an active helping paradigm. *Cognition*, *112*(2), 337–42.

Call, J., & Tomasello, M. (2008). Does the chimpanzee have a theory of mind? 30 years later. *Trends in cognitive sciences*, *12*(5), 187–92.

Carlson, S., & Moses, L. (2001). Individual differences in inhibitory control and children's theory of mind. *Child development*, *72*(4), 1032–1053.

Carston, R. (2002). *Thoughts and Utterances: The Pragmatics of Explicit Communication*. Oxford: Blackwell.

Carston, R. (2010). Lexical pragmatics, ad hoc concepts and metaphor: from a relevance theory perspective. *Italian Journal of Linguistics*, *22*(1), 153–180.

Chambers, C., & Juan, V. (2008). Perception and presupposition in real-time language comprehension: Insights from anticipatory processing. *Cognition*, *108*(1), 26–50.

Charman, T., & Baron-Cohen, S. (1992). Understanding Drawings and Beliefs: a Further Test of the Metarepresentation Theory of Autism: a Research Note. *Journal of child Psychology and Psychiatry*, *33*(6), 1105–1112.

Chemla, E., & Schlenker, P. (2012). Incremental vs. symmetric accounts of presupposition projection: An experimental approach. *Natural language semantics*, *20*(2), 177–226.

Chevallier, C. (2012). Theory of mind and autism: Beyond Baron-Cohen et al.'s Sally-Anne study. In A. Slater & P. Quinn (Eds.), *Refreshing developmental psychology: Beyond the classic studies* (pp. 148–163). Sage.

Chevallier, C., Kohls, G., Troiani, V., Brodkin, E. S., & Schultz, R. T. (2012). The social motivation theory of autism. *Trends in cognitive science*, *16*(4), 231–239.

Chevallier, C., Noveck, I. A., Happé, F., & Deirdre, W. (2011). What's in a voice? Prosody as a test case for the Theory of Mind account of autism. *Neuropsychologia*, *49*(3), 507–517.

Chevallier, C., & Wilson, D. (2010). Scalar inferences in autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *40*(9), 1104–1117.

Ciaramidaro, A., Adenzato, M., Enrici, I., Erk, S., Pia, L., Bara, B. G., & Walter, H. (2007). The intentional network: how the brain reads varieties of intentions. *Neuropsychologia*, *45*(13), 3105–3113.

Clark, H H, & Gerrig, R. J. (1984). On the pretense theory of irony. *Journal of experimental psychology. General*, *113*(1), 121–6.

Clark, H., & Carlson, T. (1982). Hearers and speech acts. *Language*, *58*(2), 332–373.

Clark, Herbert H, & Gerrig, R. J. (1984). On the Pretense Theory of Irony, *113*(1), 121–126.

Colston, H. L. (2001). On necessary conditions for verbal irony comprehension. *Pragmatics & Cognition*, *8*(2), 277–324.

Colston, H. L. (2002). Contrast and assimilation in verbal irony. *Journal of Pragmatics*, *34*, 111–142.

Colston, H. L., & O'Brien, J. (2000). Contrast and pragmatics in figurative language: Anything understatement can do, irony can do better. *Journal of Pragmatics*, *32*(11), 1557–1583.

Cornejol, C., Simonetti, F., Aldunate, N., Ibáñez, A., López, V., & Melloni, L. (2007). Electrophysiological evidence of different interpretative strategies in irony comprehension. *Journal of psycholinguistic research*, *36*(6), 411–30.

Coulson, S., & Lovett, C. (2010). Comprehension of non-conventional indirect requests: An event-related brain potential study. *Italian Journal of Linguistics*, *22*(1), 107–124.

Cummins, C., Sauerland, U., & Solt, S. (2012). Granularity and scalar implicature in numerical expressions. *Linguistics and Philosophy*, *35*(2), 135–169.

Currie, G. (2006). Why irony is pretence. In S. Nichols (Ed.), *The Architecture of the Imagination* (pp. 111–133). Oxford University Press.

Davies, C., & Katsos, N. (2010). Over-informative children: Production/comprehension asymmetry or tolerance to pragmatic violations? *Lingua*, *120*(8), 1956–1972.

Deamer, F., Pouscoulous, N., & Breheny, R. (2010). A contrastive look at metaphor and hyperbole. *UCL Working Papers in Linguistics*, *22*.

Dement, W., & Kleitman, N. (1957). Cyclic variations in EEG during sleep and their relation to eye movements, body motility, and dreaming. *Electroencephalography and Clinical Neurophysiology*, *9*(4), 673–690.

Dennett, D. (1987). *The Intentional Stance*. Cambridge, Massachusetts: MIT Press.

Dennett, D. (2009). Intentional System Theory. In B. McLaughlin, A. Beckermann, & S. Walter (Eds.), *The Oxford Handbook of Philosophy of Mind* (pp. 339–350). OUP.

Dews, S., & Winner, E. (1999). Obligatory processing of literal and nonliteral meanings in verbal irony. *Journal of pragmatics*, *31*, 1579–1599.

Downar, J., Crawley, A., Mikulis, D., & Davis, K. (2000). A multimodal cortical network for the detection of changes in the sensory environment. *Nature Neuroscience*, *3*, 277–283.

Enrici, I., Adenzato, M., Cappa, S., Bara, B. G., & Tettamanti, M. (2011). Intention processing in communication: a common brain network for language and gestures. *Journal of cognitive neuroscience*, *23*(9), 2415–31.

Eviatar, Z., & Just, M. A. (2006). Brain correlates of discourse processing: An fMRI investigation of irony and conventional metaphor comprehension. *Neuropsychologia*, *44*(12), 2348–2359.

Faust, M., & Mashal, N. (2007). The role of the right cerebral hemisphere in processing novel metaphoric expressions taken from poetry: A divided visual field study. *Neuropsychologia*, *45*(4), 860–870.

Ferstl, E.C., Neumann, J., Bogler, C., & Von Cramon, D. Y. (2008). The extended language network: A meta-analysis of neuroimaging studies on text comprehension. *Human brain mapping*, *29*(5), 581.

Ferstl, Evelyn C., & von Cramon, D. Y. (2002). What Does the Frontomedian Cortex Contribute to Language Processing: Coherence or Theory of Mind? *NeuroImage*, *17*(3), 1599–1612.

Filik, R., & Moxey, L. (2010). The on-line processing of written irony. *Cognition*, *116*(3), 421–436.

Filippova, E., & Astington, J. W. (2008). Further Development in Social Reasoning Revealed in Discourse Irony Understanding. *Memory*, *79*(1), 126 – 138.

Fischler, I., & Bradley, M. (2006). Event-related potential studies of language and emotion: words, phrases, and task effects. *Progress in brain research: Understanding Emotions*, *156*.

Friederici, A., & Meyer, M. (2004). The brain knows the difference: Two types of grammatical violations. *Brain research*, *1000*(1-2), 72–77.

Friston, K. J., Buechel, C., Fink, G., Morris, J., Rolls, E., & Dolan, R. (1997). Psychophysiological and Modulatory Interactions in Neuroimaging. *Neuroimage*, *6*(3), 218–229.

Frith, C. D., & Frith, U. (2006). The Neural Basis of Mentalizing Minireview. *NeuroImage*, 531–534.

Frith, U., & Happé, F. (1994). Autism: beyond "theory of mind." *Cognition*, *50*(1-3), 115–32.

Gallagher, H. L., Happé, F., Brunswick, N., Fletcher, P. C., Frith, U., & Frith, C. D. (2000). Reading the mind in cartoons and stories: an fMRI study of "theory of mind" in verbal and nonverbal tasks. *Neuropsychologia*, *38*(1), 11–21.

Geangu, E., Gibson, A., Kaduk, K., & Reid, V. M. (2012). The neural correlates of passively viewed sequences of true and false beliefs. *Social Cognitive and Affective Neuroscience*.

Genovese, C. R., Lazar, N. A., & Nichols, T. (2002). Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *NeuroImage*, *15*(4), 870–878.

Gernsbacher, M. A., & Robertson, D. A. (2004). Watching the brain comprehend discourse. In A. Healy (Ed.), *Experimental cognitive psychology and its applications* (pp. 157–67). Washington, DC: APA Publications.

Gerrig, R. J., & Goldvarg, Y. (2000). Additive effects in the perception of sarcasm: Situational disparity and echoic mention. *Metaphor & Symbol*, *15*, 197–208.

Gibbs, R. W. (1986). On the psycholinguistics of sarcasm. *Journal of Experimental Psychology: General*, *115*(1), 3–15.

Gibbs, R. W. (1994). *The poetics of mind: Figurative thought, language, and understanding*. Cambridge: Cambridge University Press.

Gibbs, R. W. (1999). Interpreting what speakers say and implicate. *Brain and Language*, *68*(3), 466–485.

Gibbs, R. W. (2000). Irony in talk among friends. *Metaphor and Symbol*, *15*(1-2), 5–27.

Gibbs, R. W. (2002). A new look at literal meaning in understanding what is said and implicated. *Journal of Pragmatics*, *34*(4), 457–486.

Gibbs, R. W., & Moise, J. (1997). Pragmatics in understanding what is said. *Cognition*, *62*, 51–74.

Giora, R. (1995). On Irony and Negation. *Discourse Processes*, *19*, 239–264.

Giora, R. (1997). Understanding figurative and literal language: The graded salience hypothesis. *Cognitive Linguistics*, *8*(3), 183–206.

Giora, R. (2002). Literal vs. figurative language: Different or equal? *Journal of Pragmatics*, *34*(4), 487–506.

Giora, R. (2003). *On our mind: Salience, context, and figurative language*. New York: Oxford University Press.

Giora, R. (2011). Will anticipating irony facilitate it immediately? In M. Dynel (Ed.), *The Pragmatics of Humour Across Discourse Domains* (pp. 19–31). Amsterdam: John Benjamins.

Giora, R., & Fein, O. (1999). Irony: Context and salience. *Metaphor and Symbol*, *14*(4), 241–257.

Giora, R., Fein, O., Kaufman, R., Eisenberg, D., & Erez, S. (2009). Does an "ironic situation" favor an ironic interpretation? In G. Brône & J. Vandaele (Eds.), *Cognitive poetics: Goals, gains and gaps* (pp. 383–399). Berlin/New York: Mouton de Gruyter.

Giora, R., Fein, O., Laadan, D., Wolfson, J., Zeituny, M., Kidron, R., Kaufman, R., et al. (2007). Expecting irony: Context versus salience-based effects. *Metaphor and Symbol*, *22*(2), 119–146.

Giora, R., Fein, O., & Schwartz, T. (1998). Irony: Graded Salience and Indirect Negation. *Metaphor and Symbol*, *13*(2), 83–101.

Giora, R., Zaidel, E., & Soroker, N. (2000). Differential effects of right-and left-hemisphere damage on understanding sarcasm and metaphor. *Metaphor and Symbol*, *15*(1986), 63–83.

Gitelman, D. R., Penny, W. D., Ashburner, J., & Friston, K. J. (2003). Modeling regional and psychophysiologic interactions in fMRI: the importance of hemodynamic deconvolution. *Neuroimage*, *19*(1), 200–207.

Glucksberg, S. (2003). The psycholinguistics of metaphor. *Trends in Cognitive Sciences*, *7*(2), 92–96.

Grauwe, S. D., Swain, A., & Holcomb, P. (2010). Electrophysiological insights into the processing of nominal metaphors. *Neuropsychologia*, *48*(7), 1965–1984.

Grice, P. (1975). Logic and Conversation. In P. Cole & Morgan J. (Eds.), *Syntax and Semantics* (pp. 41 – 58). Ney York: Academic Press.

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

Grodner, D., Klein, N., Carbary, K., & Tanenhaus, M. (2010). "Some," and possibly all, scalar inferences are not delayed: Evidence for immediate pragmatic enrichment. *Cognition*, *116*(1), 42–55.

Grossmann, T., Johnson, M. H., Lloyd-Fox, S., Blasi, A., Deligianni, F., Elwell, C., & Csibra, G. (2008). Early cortical specialization for face-to-face communication in human infants. *Proceedings of The Royal Society*, *275*(1653), 2803–11.

Grèzes, J., Frith, C., & Passingham, R. E. (2004). Inferring false beliefs from the actions of oneself and others: an fMRI study. *Neuroimage*, *21*(2), 744–750.

Hagoort, P. (2005). On Broca, brain, and binding: a new framework. *Trends in cognitive sciences*, *9*(9), 416–23.

Hagoort, P., Brown, C. M., & Osterhout, L. (1999). The neurocognition of syntactic processing. In C. Brown & P. Hagoort (Eds.), *Neurocognition of Language* (pp. 273–316). Oxford: Oxford University Press.

Hagoort, P., Hald, L., Bastiaansen, M., & Petersson, K. M. (2004). Integration of word meaning and world knowledge in language comprehension. *Science*, *304*(5669), 438–41.

Happé, F. (1993). Communicative competence and theory of mind in autism: a test of relevance theory. *Cognition*, *48*(2), 101–119.

Happé, F. (1994). An advanced test of theory of mind: Understanding of story characters' thoughts and feelings by able autistic, mentally handicapped, and normal children and adults. *Journal of autism and Developmental disorders*, *24*(2), 129–154.

Happé, F., Ronald, A., & Plomin, R. (2006). Time to give up on a single explanation for autism. *Nature neuroscience*, *9*(10), 1218–1220.

Haverkate, H. (1990). A speech act analysis of irony. *Journal of Pragmatics*, *14*(1), 77–109.

Herbert, C., Junghofer, M., & Kissler, J. (2008). Event related potentials to emotional adjectives during reading. *Psychophysiology*, *45*(3), 487–98.

Herten, M. V., Kolk, H. H. J., & Chwilla, D. J. (2005). An ERP study of P600 effects elicited by semantic anomalies. *Cognitive Brain Research*, *22*, 241 – 255.

Howell, D. C. (1997). *Statistical methods for psychology* (4th ed.). Belmont, CA: Duxbury Press.

Ivanko, S. L., & Pexman, P. M. (2003). Context Incongruity and Irony Processing Context Incongruity and Irony Processing. *Discourse Processes*, *35*(3), 241 – 279.

Jenkins, A. C., & Mitchell, J. P. (2010). Mentalizing under uncertainty: dissociated neural responses to ambiguous and unambiguous mental state inferences. *Cerebral cortex*, *20*(2), 404–410.

Jorgensen, J., Miller, G., & Sperber, D. (1984). Test of the mention theory of irony. *Journal of Experimental Psychology. General*, *113*(1), 112–120.

Josephs, O., Turner, R., & Friston, K. J. (1997). Event-related fMRI. *Human brain mapping*, *5*(4), 243–248.

Jung-Beeman, M. (2005). Bilateral brain processes for comprehending natural language. *Trends in cognitive sciences*, *9*(11), 512–518.

Kacinik, N., & Chiarello, C. (2007). Understanding metaphors: Is the right hemisphere uniquely involved? *Brain and Language*, *100*(2), 188–207.

Kalbe, E., Schlegel, M., Sack, A. T., Nowak, D. A., Dafotakis, M., Bangard, C., Brand, M., et al. (2010). Dissociating cognitive from affective theory of mind: a TMS study. *Cortex*, *46*(6), 769–780.

Kampe, K. K. W., Frith, C. D., & Frith, U. (2003). "Hey John": signals conveying communicative intention toward the self activate brain regions associated with "mentalizing," regardless of modality. *The Journal of neuroscience*, *23*(12), 5258–5263.

Kanske, P., & Kotz, S. a. (2007). Concreteness in emotional words: ERP evidence from a hemifield study. *Brain research*, *1148*, 138–48.

Kaplan, J., Brownell, H., Jacobs, J., & Gardner, H. (1990). The effects of right hemisphere damage on the pragmatic interpretation of conversational remarks. *Brain and Language*, *38*(2), 315–333.

Katz, A., Blasko, D., & Kazmerski, V. (2004). Saying what you don't mean social influences on sarcastic language processing. *Current Directions in Psycological Science*, *13*(5), 186–189.

Katz, A., & Lee, C. (1993). The Role of Authorial Intent in Determining Verbal Irony and Metaphor. *Metaphor and Symbol*, *8*(4), 257–279.

Katz, A., & Pexman, P. (1997). Interpreting Figurative Statements: Speaker Occupation Can Change Methaphor to Irony. *Metaphor and Symbol*, *12*(1), 19–41.

Keenan, T., & Quigley, K. (1999). Do young children use echoic information in their comprehension of sarcastic speech? A test of echoic mention theory. *British Journal of Developmental Psychology*, *17*, 83–96.

Keysar, B, Barr, D., Balin, J., & Brauner, J. (2000). Taking perspective in conversation: The role of mutual knowledge in comprehension. *Psychological Science*, *11*(1), 32–38.

Keysar, Boaz, Lin, S., & Barr, D. J. (2003). Limits on theory of mind use in adults. *Cognition*, *89*(1), 25–41.

Kissler, J., Herbert, C., Winkler, I., & Junghofer, M. (2009). Emotion and attention in visual word processing: an ERP study. *Biological Psychology*, *80*(1), 75–83.

Klimesch, W. (1994). *The structure of long-term memory: A connectivity model of semantic processing*. Hillsdale, NJ: Lawrence Erlbaum.

Klimesch, W. (1996). Memory processes, brain oscillations and EEG synchronization. *International Journal of Psychophysiology*, *24*(1-2), 61–100.

Klimesch, W. (1999). EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. *Brain research reviews*, *29*(2-3), 169–195.

Klimesch, W., Doppelmayr, M., Pachinger, T., & Ripper, B. (1997). Brain oscillations and human memory: EEG correlates in the upper alpha and theta band. *Neuroscience letters*, *238*(1-2), 9–12.

Klimesch, W., Doppelmayr, M., Schimke, H., & Ripper, B. (1997). Theta synchronization and alpha desynchronization in a memory task. *Psychophysiology*, *34*(2), 169–176.

Klimesch, W., Doppelmayr, M., Schwaiger, J., Auinger, P., & Winkler, T. (1999). "Paradoxical" alpha synchronization in a memory task. *Cognitive brain research*, *7*(4), 493–501.

Kreuz, R. J., Kassler, M. A., Coppenrath, L., & McLain Allen, B. (1999). Tag questions and common ground effects in the perception of irony. *Journal of Pragmatics*, *31*, 1685–1700.

Kreuz, R. J., & Link, K. E. (2002). Asymmetries in the use of verbal irony. *Journal of Language and Social Psychology*, *21*, 127–143.

Kreuz, R. J., Roberts, R. M., Johnson, B. K., & Bertus, L. E. (1996). Figurative language occurrence and co-occurence in contemporary literature. In R. J. Kreuz & M. S. MacNealy (Eds.), *Empirical approaches to literature and aesthetics* (pp. 83–97). Norwood, NJ: Ablex.

Kumon-Nakamura, S., Glucksberg, S., & Brown, M. (1995). How about another piece of pie: the allusional pretense theory of discourse irony. *Journal of experimental psychology. General*, *124*(1), 3–21.

Kuperberg, G. (2007). Neural mechanisms of language comprehension: Challenges to syntax. *Brain Research*, *1146*, 23–49.

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, *207*(4427), 203–205.

Lancker, D. V., & Kempler, D. (1987). Comprehension of familiar phrases by left-but not by right-hemisphere damaged patients. *Brain and language*, *32*(2), 265–277.

Langdon, R., Davies, M., & Coltheart, M. (2002). Understanding minds and understanding communicated meanings in schizophrenia. *Mind & Language*, *17*(1-2), 68–104.

Lattner, S., & Friederici, A. D. (2003). Talker's voice and gender stereotype in human auditory sentence processing – evidence from event-related brain potentials. *Neuroscience Letters*, *339*(3), 191–194.

Lee, S., & Dapretto, M. (2006). Metaphorical vs. literal word meanings: fMRI evidence against a selective role of the right hemisphere. *Neuroimage*, *29*(2), 536–544.

Leslie, A M, & Thaiss, L. (1992). Domain specificity in conceptual development: neuropsychological evidence from autism. *Cognition*, *43*(3), 225–251.

Leslie, Alan M. (1987). Pretense and representation: The origins of "theory of mind." *Psychological Review*, *94*(4), 412–426.

Lev-Rana, S., Shamay-Tsoory, S. G., Zangenc, A., & Levkovitz, Y. (2012). Transcranial magnetic stimulation of the ventromedial prefrontal cortex impairs theory of mind learning. *European Psychiatry*, *27*(4), 285–289.

Lewis, C., & Osborne, A. (1990). Three‑Year‑Olds' Problems with False Belief: Conceptual Deficit or Linguistic Artifact? *Child Development*, *61*(5), 1514–1519.

Liu, D., Sabbagh, C. A., Gehring, W. J., & Wellman, H. M. (2004). Decoupling beliefs from reality in the brain: an ERP study of theory of mind. *NeuroReport*, *15*(6), 991–995.

Long, D. L., & Baynes, K. (2002). Discourse representation in the two cerebral hemispheres. *Journal of cognitive neuroscience*, *14*(2), 228–242.

Ma, N., Vandekerckhove, M., Van Overwalle, F., Seurinck, R., & Fias, W. (2011). Spontaneous and intentional trait inferences recruit a common mentalizing network to a different degree: spontaneous inferences activate only its core areas. *Social neuroscience*, *6*(2), 123–138.

Malle, B. (2002). The relation between language and theory of mind in development and evolution. In T. Givon & B. F. Malle (Eds.), *The evolution of language out of pre-language* (pp. 265–284). Amsterdam: John Benjamins B. V.

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. New York: W.H. Freeman.

Mashal, N., Faust, M., Hendler, T., & Jung-Beeman, M. (2009). An fMRI study of processing novel metaphoric sentences. *Laterality: Asymmetries of Body, Brain and Cognition*, *14*(1), 30–54.

McDonald, S. (1999). Exploring the process of inference generation in sarcasm: a review of normal and clinical studies. *Brain and language*, *68*(3), 486–506.

Meinhardt, J, Sodian, B., & Thoermer, C. (2011). True-and false-belief reasoning in children and adults: An event-related potential study of theory of mind. *Developmental Cognitive Neuroscience*, *1*(1), 67–76.

Meinhardt, Jörg, Kühn-Popp, N., Sommer, M., & Sodian, B. (2012). Distinct neural correlates underlying pretense and false belief reasoning: Evidence from ERPs. *NeuroImage*, *63*(2), 623–631.

Mitchell, J. P. (2008). Activity in right temporo-parietal junction is not selective for theory-of-mind. *Cerebral cortex*, *18*(2), 262–271.

Mitchell, J. P. (2009). Inferences about mental states. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, *364*(1521), 1309–16.

Mitchell, J. P., Macrae, C. N., & Banaji, M. R. (2006). Dissociable medial prefrontal contributions to judgments of similar and dissimilar others. *Neuron*, *50*(4), 655–63.

Moxey, L. (2006). Effects of what is expected on the focussing properties of quantifiers: A test of the presupposition-denial account. *Journal of Memory and Language*, *55*(3), 422–439.

Nadig, A. S., & Sedivy, J. C. (2002). Evidence of Perspective-Taking Constraints in Children's On-Line Reference Resolution. *Psychological Science*, *13*(4), 329–336.

Nieuwland, M. S., & Berkum, J. J. A. V. (2005). Testing the limits of the semantic illusion phenomenon : ERPs reveal temporary semantic change deafness in discourse comprehension. *Cognitive Brain Research*, *24*(3), 691 – 701.

Nieuwland, M. S., 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*(3), 324–346.

Nieuwland, M. S., & Van Berkum, J. (2006). When peanuts fall in love: N400 evidence for the power of discourse. *Journal of Cognitive Neuroscience*, *18*(7), 1098–1111.

Noveck, I. A. (2001). When children are more logical than adults: Experimental investigations of scalar implicature. *Cognition*, *78*(2), 165–188.

Noveck, I. A., & Spotorno, N. Narrowing, to appear in L. Goldstein (Ed.), *Brevity*.

Noveck, I. A., & Reboul, A. (2008). Experimental pragmatics: a Gricean turn in the study of language. *Trends in Cognitive Sciences*, *12*(11), 425–431.

Noveck, I. A., & Sperber, D. (2004). *Experimental Pragmatics*. London: Palgrave.

Noveck, I. A., & Posada, A. (2003). Characterizing the time course of an implicature: An evoked potentials study. *Brain and Language*, *85*(2), 203–210.

Nurmsoo, E., & Bloom, P. (2008). Preschoolers' Perspective Taking in Word Learning Do They Blindly Follow Eye Gaze? *Psychological science*, *9*(3), 211–215.

Onishi, K., & Baillargeon, R. (2005). Do 15-Month-Old Infants Understand False Beliefs? *Science*, *308*(5719), 255–258.

Ozonoff, S., Pennington, B. F., & Rogers, S. J. (1991). Executive function deficits in high-functioning autistic Individuals: Relationship to Theory of Mind. *Journal of child Psychology and Psychiatry*, *32*(7), 1081–1105.

Peleg, O., Giora, R., & Fein, O. (2001). Salience and Context Effects: Two Are Better Than One. *Metaphor and Symbol*, *16*(3-4), 173–192.

Pennington, B. F. (2006). From single to multiple deficit models of developmental disorders. *Cognition*, *101*(2), 385–413.

Perner, J., Aichhorn, M., Kronbichler, M., Staffen, W., & Ladurner, G. (2006). Thinking of mental and other representations: The roles of left and right temporo-parietal junction. *Social Neuroscience*, *1*(3-4), 245–258.

Peterson, C., & Siegal, M. (1995). Deafness, conversation and theory of mind. *Journal of child Psychology and Psychiatry*, *36*(3), 459–474.

Peterson, C., & Siegal, M. (1999). Representing Inner Worlds: Theory of Mind in Autistic , Deaf , and Normal Hearing Children. *Psychological Science*, *10*(2), 126–129.

Pexman, P. M., & Zvaigzne, M. T. (2004). Does irony go better with friends? *Metaphor & Symbol*, *19*, 143–163.

Pexman, P., & Olineck, K. (2002). Understanding Irony How Do Stereotypes Cue Speaker Intent? *Journal of Language and Social Psychology*, *21*(3), 245–274.

Pexman, P., Whalen, J., & Green, J. (2010). Understanding Verbal Irony: Clues From Interpretation of Direct and Indirect Ironic Remarks. *Discourse Processes*, *47*(3), 237–261.

Pfurtscheller, G., & Lopes da Silva, F. H. E. (1999). Event-related desynchronization. *Handbook of Electroencephalography and Clinical Neurophysiology, revised series (Vol. 6)*. Amsterdam: Elsevier.

Pfurtscheller, G., Stancak, A., & Neuper, C. (1996). Event-related synchronization (ERS) in the alpha band—an electrophysiological correlate of cortical idling: a review. *International Journal of Psychophysiology*, *24*(1-2), 39–46.

Pobric, G., Mashal, N., Faust, M., & Lavidor, M. (2008). The role of the right cerebral hemisphere in processing novel metaphoric expressions: A transcranial magnetic stimulation study. *Journal of Cognitive Neuroscience*, *20*(1), 170–181.

Pouscoulous, N., Noveck, I. A., Politzer, G., & Bastide, A. (2007). A developmental investigation of processing costs in implicature production. *Language Acquisition*, *14*(4), 347–375.

Premack, D. (1983). The codes of man and beasts. *Behavioral and Brain Sciences*, *6*(1), 125–136.

Premack, D., & Woodruff, G. (1978a). Does the chimpanzee have a Theory of Mind? *Behavioral and brain sciences*, *1*(4), 515–526.

Premack, D., & Woodruff, G. (1978b). Does the chimpanzee have a theory of mind? *Behavioral and brain sciences*, *1*(4), 515–526.

Rapp, A., Leube, D., Erb, M., Grodd, W., & Kircher, T. (2007). Laterality in metaphor processing: Lack of evidence from functional magnetic resonance imaging for the right hemisphere theory. *Brain and Language*, *100*(2), 142–149.

Rapp, A. M., Mutschler, D., Wild, B., Erb, M., Lengsfeld, I., Saur, R., & Grodd, W. (2010). Neural correlates of irony comprehension: The role of schizotypal personality traits. *Brain and language*, *113*(1), 1–12.

Rapp, A.M., Leube, D. T., Erb, M., Grodd, W., & Kircher, T. T. (2004). Neural correlates of metaphor processing. *Brain Research*, *20*(3), 395–402.

Rapp, Alexander M, Mutschler, D. E., & Erb, M. (2012). Where in the brain is nonliteral language? A coordinate-based meta-analysis of functional magnetic resonance imaging studies. *NeuroImage*, *63*(1), 600–610.

Recanati, F. (2007). Indexicality, context and pretence. In N. Burton-Roberts (Ed.), *Pragmatics* (pp. 213–229). Basingstoke: Palgrave Macmillan.

Regel, S., Coulson, S., & Gunter, T. C. (2010). The communicative style of a speaker can affect language comprehension? ERP evidence from the comprehension of irony. *Brain research*, *1311*, 121–35.

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*(2), 277–293.

Roberts, R. M., & Kreuz, R. J. (1994). Why do people use figurative language? *Psychological Science*, *5*, 159–163.

Roth, D., & Leslie, A. M. (1998). Solving belief problems: toward a task analysis. *Cognition*, *66*(1), 1–31.

Rubio Fernández, P. (2007). Suppression in metaphor interpretation: Differences between meaning selection and meaning construction. *Journal of semantics*, *24*(4), 345–371.

Ruffman, T., & Perner, J. (2005). Do infants really understand false belief? Respense to Leslie. *Trends in cognitive sciences*, *9*(10), 462–463.

Röhm, D., Klimesch, W., Haider, H., & Doppelmayr, M. (2001). The role of theta and alpha oscillations for language comprehension in the human electroencephalogram. *Neuroscience Letters*, *310*(2-3), 137–140.

Sabbagh, M. A., Xu, F., Carlson, S. M., Moses, L. J., & Lee, K. (2006). The Development of Executive Functioning and Theory of Mind: A Comparison of Chinese and U.S. Preschoolers. *Psychological Science*, *17*(1), 74–81.

Sabbagh, M., & Taylor, M. (2000). NEURAL CORRELATES OF THEORY-OF-MIND REASONING: An Event-Related Potential Study. *Psychological Science*, *11*(1), 46–50.

Saxe, R. (2006). Uniquely human social cognition. *Current opinion in neurobiology*, *16*(2), 235–9.

Saxe, R. (2010). The right temporo-parietal junction: a specific brain region for thinking about thoughts. *Handbook of Theory of Mind* (pp. 1–35).

Saxe, R., Brett, M., & Kanwisher, N. (2006). Divide and conquer: a defense of functional localizers. *NeuroImage*, *30*(4), 1088–96.

Saxe, R., Carey, S., & Kanwisher, N. (2004). Understanding other minds: linking developmental psychology and functional neuroimaging. *Annual review of psychology*, *55*, 87–124.

Saxe, R., & Kanwisher, N. (2003). People thinking about thinking people: The role of the temporo-parietal junction in "theory of mind." *NeuroImage*, *19*(4), 1835–1842.

Saxe, R., & Powell, L. J. (2006). It's the thought that counts: specific brain regions for one component of theory of mind. *Psychological science*, *17*(8), 692–9.

Saxe, R., & Wexler, A. (2005). Making sense of another mind: the role of the right temporo-parietal junction. *Neuropsychologia*, *43*(10), 1391–9.

Schmidt, G., DeBuse, C., & Seger, C. (2007). Right hemisphere metaphor processing? Characterizing the lateralization of semantic processes. *Brain and language*, *100*(2), 127–141.

Scholz, J., Triantafyllou, C., Whitfield-Gabrieli, S., Brown, E. N., & Saxe, R. (2009). Distinct regions of right temporo-parietal junction are selective for theory of mind and exogenous attention. *PloS one*, *4*(3), e4869.

Schwoebel, J., Dews, S., Winner, E., & Srinivas, K. (2000). Obligatory processing of the literal meaning of ironic utterances: further evidence. *Metaphor and Symbol*, *15*(1 & 2), 47–61.

Searle, J. R. (1969). *Speech Acts. An Essay in the Philosophy of Language*. Cambridge: Cambridge University Press.

Searle, J. R. (1979). *Expression and meaning. Studies in the theory of speech acts*. Cambridge: Cambridge University Press.

Sedivy, J. C. (2003). Pragmatic versus form-based accounts of referential contrast: evidence for effects of informativity expectations. *Journal of psycholinguistic research*, *32*(1), 3–23.

Shamay-Tsoory, S. G., & Aharon-Peretz, J. (2007). Dissociable prefrontal networks for cognitive and affective theory of mind: a lesion study. *Neuropsychologia*, *45*(13), 3054–3067.

Shamay-Tsoory, S. G., Aharon-Peretz, J., & Levkovitz, Y. (2007). The neuroanatomical basis of affective mentalizing in schizophrenia: Comparison of patients with schizophrenia and patients with localized prefrontal lesions. *Schizophrenia Research*, *90*(1-3), 274–283.

Shamay-Tsoory, S. G., Tomer, R., & Aharon-Peretz, J. (2005). The neuroanatomical basis of understanding sarcasm and its relationship to social cognition. *Neuropsychology*, *19*(3), 288–300.

Shibata, M., Toyomura, A., Itoh, H., & Abe, J. (2010). Neural substrates of irony comprehension: A functional MRI study. *Brain research*, *1308*, 114–123.

Siegal, M., & Varley, R. (2006). Aphasia, language, and theory of mind. *Social neuroscience*, *1*(3-4), 167–174.

Singer, W. (1993). Synchronization of cortical activity. *Annual Rev of Psychophysiologyv of Psychophysiology*, *55*, 349–374.

Singer, Wolf. (1999). Neuronal Synchrony : A Versatile Code for the Definition of Relations ? Most of our knowledge about the functional organization. *Neuron*, *24*(1), 49–65.

Slaughter, V. (1998). Children's understanding of pictorial and mental representations. *Child Development*, *69*(2), 321–332.

Slugoski, B. R., & Turnbull, W. (1988). Cruel to be kind and kind to be cruel: Sarcasm, banter, and social relations. *Journal of Language and Social Psychology*, *7*, 101–121.

Sotillo, M., Carretié, L., Hinojosa, J. a, Tapia, M., Mercado, F., López-Martín, S., & Albert, J. (2005). Neural activity associated with metaphor comprehension: spatial analysis. *Neuroscience letters*, *373*(1), 5–9.

Southgate, V., Chevallier, C., & Csibra, G. (2010). Seventeen-month-olds appeal to false beliefs to interpret others' referential communication. *Developmental Science*, *13*(6), 907–912.

Speer, N. K., Zacks, J. M., & Reynolds, J. R. (2007). Human brain activity time-locked to narrative event boundaries. *Psychological science*, *18*(5), 449–455.

Sperber, D., & Wilson, D. (1981). Irony and the use-mention distinction. In P. Cole (Ed.), *Radical pragmatics* (pp. 295–318). New York: Academic Press Elsevier.

Sperber, D., & Wilson, D. (1986). *Relevance: Communication and Cognition*. Oxford: Blackwell; Wiley-Blackwell; 2nd Edition edition 1995.

Sperber, D., & Wilson, D. (2002). Pragmatics, Modularity and Mind-reading. *Mind and Language*, *17*(1-2), 3–23.

Sperber, D., & Wilson, D. (2008). A deflationary account of metaphors. In R. W. Gibbs (Ed.), *The Cambridge handbook of metaphor and thought* (pp. 84–105). New York: Cambridge Univesity Press.

Spotorno, N., Koun, E., Prado, J., Van Der Henst, J.-B., & Noveck, I. A. (2012). Neural evidence that utterance-processing entails mentalizing: The case of irony. *NeuroImage*, *63*(1), 25–39.

Sterman, M., Kaiser, D., & Veigel, B. (1996). Spectral Analysis of Event-Related EEG Responses During Short-Term Memory Performance. *Brain topography*, *9*(1), 21–30.

Stone, V. E., Baron-Cohen, S., Calder, A., Keane, J., & Young, A. (2003). Acquired theory of mind impairments in individuals with bilateral amygdala lesions. *Neuropsychologia*, *41*(2), 209–20.

Stringaris, A. K., Medford, N., Giora, R., Giampietro, V. C., Brammer, M. J., & David, A. S. (2006). How metaphors influence semantic relatedness judgments: the role of the right frontal cortex. *NeuroImage*, *33*(2), 784–93.

Surian, L., Caldi, S., & Sperber, D. (2007). Attribution of beliefs by 13-month-old infants. *Psychological science*, *18*(7), 580–6.

Tager-Flusberg, H. (2000). Language and Understanding Minds in Autism. In S. Baron-Cohen, H. Tager-Flusberg, & D. J. Cohen (Eds.), *Understanding other minds: Perspectives from autism and developmental cognitive neuroscience* (pp. 124–149). Oxford: Oxford University Press.

Talairach, J., & Tournoux, P. (1988). *Co-planar stereotaxic atlas of the human brain: 3-dimensional proportional system: an approach to cerebral imaging*. New York: Thieme.

Tallon-Baudry, C., Bertrand, O., Delpuech, C., & Permier, J. (1997). Oscillatory gamma-band (30-70 Hz) activity induced by a visual search task in humans. *The Journal of neuroscience*, *17*(2), 722–734.

Tompkins, C., & Mateer, C. (1985). Right hemisphere appreciation of prosodic and linguistic indications of implicit attitude. *Brain and language*, *24*(2), 185–203.

Toplak, M., & Katz, A. (2000). On the uses of sarcastic irony. *Journal of pragmatics*, *32*, 1467–1488.

Uchiyama, H. T., Saito, D. N., Tanabe, H. C., & Harada, T. (2011). Distinction between the literal and intended meanings of sentences: A functional magnetic resonance imaging study of metaphor and sarcasm. *Cortex*, *48*(5), 563–583.

Uchiyama, H. T., Seki, A., Kageyama, H., Saito, D. N., Koeda, T., Ohno, K., & Sadato, N. (2006). Neural substrates of sarcasm: a functional magnetic-resonance imaging study. *Brain research*, *1124*(1), 100–110.

Utsumi, A. (2000). Verbal irony as implicit display of ironic environment: Distinguishing ironic utterances from non irony. *Journal of Pragmatics*, *32*, 1777–1806.

Van Berkum, J. (2004). Sentence comprehension in a wider discourse: Can we use ERPs to keep track of things. In M. Carreiras & C. Clifton (Eds.), *The on-line study of sentence*

comprehension: eyetracking, ERPs and beyond (pp. 229–270). New York: Psychology Press.

Van Overwalle, F. (2009). Social cognition and the brain: a meta-analysis. *Human brain mapping*, *30*(3), 829–58.

Van Overwalle, F., & Baetens, K. (2009). Understanding others' actions and goals by mirror and mentalizing systems: a meta-analysis. *NeuroImage*, *48*(3), 564–84.

Varela, F., & Lachaux, J. (2001). The brainweb: phase synchronization and large-scale integration. *Nature reviews Neuroscience*, *2*(April).

Vossel, S., Weidner, R., Thiel, C., & Fink, G. (2009). What is "odd" in Posner's location-cueing paradigm? Neural responses to unexpected location and feature changes compared. *Journal of Cognitive Neuroscience*, *21*(1), 30–41.

Wakusawa, K., Sugiura, M., Sassa, Y., Jeong, H., Horie, K., Sato, S., Yokoyama, H., et al. (2007). Comprehension of implicit meanings in social situations involving irony: A functional MRI study. *NeuroImage*, *37*(4), 1417–1426.

Walton, K. (1990). *Mimesis as Make-believe: On the Foundations of the Representational Arts*. Cambridge, Massachusetts: Harvard University Press.

Wang, A. T., Lee, S. S., Sigman, M., & Dapretto, M. (2006). Developmental changes in the neural basis of interpreting communicative intent. *Social cognitive and affective neuroscience*, *1*(2), 107–121.

Weiss, S., Rappelsberger, P., Schack, B., & Mueller, H. M. (2003). Koharenz- und Phasenuntersuchungen und ihre Bedeutung fur die Untersuchung von Sprachprozessen. In G. R. H. M. Mueller (Ed.), *Neurokognition der Sprache*. Tubingen: Stauffenburg.

Weiss, Sabine, Mueller, H. M., Schack, B., King, J. W., Kutas, M., & Rappelsberger, P. (2005). Increased neuronal communication accompanying sentence comprehension. *International journal of psychophysiology*, *57*(2), 129–41.

Wellman, H M, Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: the truth about false belief. *Child development*, *72*(3), 655–684.

Wellman, Henry M, & Liu, D. (2004). Scaling of theory-of-mind tasks. *Child development*, *75*(2), 523–41.

Wilson, D. (2006). The pragmatics of verbal irony: Echo or pretence? *Lingua*, *116*(10), 1722–1743.

Wilson, D. (2009). Irony and metarepresentation. *UCLWPL*, *21*, 183–226.

Wilson, D., & Carston, R. (2006). Metaphor, relevance and the "emergent property"issue. *Mind and Language*, *21*(3), 403–433.

Wilson, D., & Sperber, D. (1992). On verbal irony. *Lingua*, *87*(1), 53–76.

Wilson, D., & Sperber, D. (2004). Relevance Theory. In L. R. Horn & G. Ward (Eds.), *The Handbook of Pragmatics* (pp. 607–632). Oxford: Blackwell.

Wilson, D., & Sperber, D. (2012a). Explaining Irony. In D. Wilson & D. Sperber (Eds.), *Meaning and Relevance*. Cambridge: Cambridge University Press.

Wilson, Deirdre, & Sperber, D. (2012b). *Meaning and Relevance*. Cambridge: Cambridge University Press.

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, *13*(1), 103–128.

Winner, E., Brownell, H., Happé, F., Blum, A., & Pincus, D. (1998). Distinguishing lies from jokes: Theory of Mind deficits and discourse interpretation in right hemisphere brain-damaged patients. *Brain and language*, *62*(1), 89–106.

Winner, E., & Gardner, H. (1977). The comprehension of metaphor in brain-damaged patients. *Brain*, *100*(4), 717–729.

Yazdi, A. A., German, T. P., Defeyter, M. A., & Siegal, M. (2006). Competence and performance in belief-desire reasoning across two cultures: the truth, the whole truth and nothing but the truth about false belief? *Cognition*, *100*(2), 343–68.

Zaitchik, D. (1990). When representations conflict with reality: The preschooler's problem with false beliefs and "false" photographs. *Cognition*, *35*(1), 41–68.

de Villiers, P. A. (2005). The role of language in theory-of-mind development: What deaf children tell us. In J. W. Astington & J. A. Baird (Eds.), *Why language matters for theory of mind* (pp. 266–297). Oxford: Oxford University Press.

van Berkum, J. J., Zwitserlood, P., Bastiaansen, M., Brown, C., & Hagoort, P. (2004). So who's "he" anyway? Differential ERP and ERSP effects of referential success, ambiguity and failure during spoken language comprehension. *Paper presented at the Annual Meeting of the Cognitive Neuroscience Society, San Francisco*.

van den Brink, D., Van Berkum, J. J. a, Bastiaansen, M. C. M., Tesink, C. M. J. Y., Kos, M., Buitelaar, J. K., & Hagoort, P. (2010). Empathy matters: ERP evidence for inter-individual differences in social language processing. *Social cognitive and affective neuroscience*, *7*(2), 173–83.