Dipartimento di Psicologia e Scienze Cognitive

Dottorato di Ricerca in Scienze Cognitive

XXXIV Ciclo

Tesi di Dottorato

Psychological Aspects of Pragmatics of Negation

Relatore

prof.ssa Chiara Finocchiaro

prof. Francesco Vespignani

Dottorando

Pietro Mingardi

anno accademico 2020-2021

Table of Contents

Abstract

The topic of this thesis is the comprehension of negative sentences. In the literature there is a number of evidence suggesting that negation is not integrated immediately into the comprehension process, but only after dealing with the affirmative counterpart meaning of the sentence (e.g. Clark & Chase, 1972; Fischler, 1983; Hasson and Glucksberg, 2006; Kaup, Yaxley, Madden, Zwaan & Luedtke, 2007). Other evidence, however, suggests that if interpreted within pragmatic felicitous contexts, the actual meaning of the negation can be immediately accessed (Nieuwland and Kuperberg, 2008; Dale & Duran, 2011; Tian, Breheny and Ferguson, 2010, 2016; Arroyo, 1982). To date, it is still unclear which exactly are the pragmatic factors that allow to access immediately the actual meaning of the negation. The objective of this thesis is to try to identify some of these possible pragmatic factors. In six different experiments, using different experimental paradigms (sentence-picture verification task, probe-recognition task, self-paced reading task) we investigated four different pragmatic factors: 1) the use of the negation to describe a minority (Wason, 1965); 2) the use of the negation to describe something that non-exists (Nordmeyer and Frank, 2015); 3) the fact that a negative sentence answers to a negative Question Under Discussion (Tian, Breheny and Ferguson, 2010); 4) the use of the negation to describe something exceptional (Arroyo, 1982). For no one of these four factors we have found clear evidence that these are key pragmatic factors that allow to access immediately the actual meaning of the negation. At the opposite, our results seem to suggest that only in really specific situations the actual meaning of the negation can be immediately accessed, in particular when pragmatic and lexical predictability factors come into play together.

Chapter 1: Introduction

Negation is difficult

Negation is a universal feature of human language (Horn, 1989) which allows us to linguistically describe the world not only as it is, but also as it is not (Zanellini, 2020). The study of negation has fascinated psychologists since a growing body of evidence has shown that such a simple and massively used word such as "not" can dramatically affect the easiness with which many psychological experimental tasks are performed.

Negation requires extra processing

Among the earliest experimental psychology investigations of the effects of negation there are some works by Paul Wason (1959, 1961). In Wason (1961) participants were asked to complete some sentences that might be affirmative or negative, such as "…. is an odd number" or "…. is not an odd number". He asked participants to complete the sentences sometimes to make them true and some other times to make them false. Analyzing the answers in the task, Wason found participants were always faster and made less mistakes to complete affirmative vs negative sentences (Wason, 1961). This result is clear evidence that participants had a greater difficulty in performing the task with negative with respect to affirmative sentences. Despite the method of Wason was criticized by Wales and Grieve (1969) and Horn (1989) in terms of lack of ecologicity, since it is unnatural to produce a negative sentence ("not odd" or "not even") when there is an affirmative description version ("even" and "odd") that is generally adopted (Zannellini, 2020), the effect of faster and more accurate responses with affirmative vs negative sentences found by Wason was also found by numerous other behavioural experiments. Many of these experiments used a task that was extensively used to study

negation, that is the truth-judgment task. In this type of task participants read a sentence and are requested to evaluate whether it is true or false, usually with respect to a picture (Gough, 1965; Gough, 1966; Slobin 1966; Clark & Chase, 1972; Clark & Chase, 1974; Carpenter & Just, 1975; Trabasso, Rollins & Shaughnessy 1971) or with respect to general world-knowledge (Wason & Jones, 1963; Wales & Grieve, 1969, Arroyo, 1982). Detrimental effects of negations have also been found in self-paced reading experiments, in which negative sentences are read at a slower pace than affirmative sentences (Glenberg, Robertson, Jansen & Johnson-Glenberg, 1999; Luedtke & Kaup, 2006), and in memory recall tasks (Cornish & Wason, 1970; MacDonald & Just, 1989; Kaup, 1997; 2001, Kaup & Zwann, 2003). For example, in a memory recall task, Cornish and Wason (1970) presented participants with a series of adjectives that described an object. These adjectives might be presented in the affirmative form (e.g. "It is bright") or in the negative form (e.g. "It is not bright"). Subsequently they asked participants to recall as many adjectives as possible. Interestingly, they found participants made significantly more errors when they had to recall negated vs affirmed adjectives. The vast majority of the errors' participants made were due to the fact they wrongly recalled a negative adjective as if it was affirmative (e.g. "not bright" was reported as "dull"). These errors were interpreted as evidence of the difficulty to keep in memory negated properties (Cornish and Wason, 1970). Similar effect to that found by Cornish and Wason seems that found by McDonalds & Just (MacDonald & Just, 1989 (Experiment 1)). In their experiment the authors presented participants with sentences such as, for example, "Elizabeth baked some bread but no cookies". After reading the sentence participants were presented with a probe word such as, for example, "bread", "cookies", "baguette" or "biscuits" and were asked whether the probe word was mentioned in the previous sentence or not. Authors found participants were slower to judge the presence of words presented after the negation. Within the above example, participants were slower to respond to "cookies" with respect to "bread".

The result was interpreted by the authors as evidence that negation reduces the activation level of the words under its scope. Independently from this interpretation, however, the result clearly shows a greater difficulty in the recall of the negated word. The result was confirmed in a second study in which participants did not have to make an explicit judgment but simply read the probe word (MacDonald & Just, 1989 (Experiment 2)), and by a series of similar experiments by Kaup (1997, 2001) and Kaup & Zwaan (2003).

Neuroimaging studies have found evidence of a greater effort in handling negative sentences with respect to affirmative ones, as indicated by larger consumption of metabolic resources by the brain (Carpenter, Just, Keller, Eddy, Thulborn; 1999; Hasegawa, Carpenter and Just, 2002; Christensen, 2009; Bahlmann, Mueller, Makuuchi & Friederici, 2011; but see Tettamanti, Manenti, Della Rosa, Falini, Perani, Cappa & Moro, 2008, Tomasino, Weiss and Fink, 2010 for some salient exceptions). For example, Carpenter and colleagues (Carpenter et al., 1999), Hasegawa and colleagues (Hasegawa et al., 2002) and Christensen (2009) investigated the BOLD signal in participants that were reading affirmative or negative sentences. Carpenter et al. (1999) found an increasing in the BOLD signal in the left temporal and in left parietal regions during negative vs affirmative sentences reading, while Hasegawa et al., (2002) and Christens (2009) found enhanced activation in the left premotor cortex. These results were interpreted by the authors as evidence of the higher computation demand negative sentences require with respect to affirmative ones. Neurophysiological experiments, instead, found evidence of the higher difficulty in handling negative vs affirmative sentences as indicated by specific electrophysiological correlates. For example, using the ERP technique, Spychalska and colleague (Spychalska, Haase, Kontinen, & Werning, 2019) found an increasing in the P600 amplitude when reading negative vs affirmative sentences. As in the fMRI studies, the result was interpreted as evidence of the more effortful processing negation requires with

respect to affirmation to be comprehended. This P600 effect for negative vs affirmative sentences was found also in an experiment by Herbert & Kübler (2011).

The Polarity by Truth-Value Interaction Effect

Until now we have seen that negation seems to be more difficult to handle than affirmation since the presence of the negation, poses problems in the execution of different psychological experimental tasks. We have also seen that evidence of this difficulty can be found in neuroimaging and neurophysiological experiments. Now, we will present another famous experimental effect that is found with negation, the so-called polarity by truth-value interaction effect. This effect is found in truth-judgment tasks and shows that whereas with affirmative sentences participants are facilitated in handling true sentences (e.g. "Joe Biden is the president of the USA") with respect to false ones (e.g. "Boris Johnson is the president of the USA"), with negative sentences it is the opposite because participants are facilitated in handling false sentences ("Joe Biden is not the president of the USA") with respect to true ones ("Boris Johnson is not the president of the USA"). The effect has been found both in behavioural and EEG experiments.

Behavioural Experiments

Probably the first evidence of the polarity by truth-value interaction effect was reported in the experiment of Wason (1961) we presented above. Indeed, in that experiment, Wason asked participants to complete some affirmative and negative sentences of the type "… is an odd number" or "…is not an odd number". As we wrote above, Wason asked participants to complete the sentences in order to make them true or false depending on a given instruction. Analyzing the reaction times and the accuracy with which participants completed the sentences, Wason found that, with affirmative sentences, participants were faster and more accurate in

completing sentences to make them true rather than false, while with negative sentences, participants were more accurate in completing the sentences to make them false rather than true. This facilitation found by Wason in handling affirmative true and negative false sentences with respect to affirmative false and negative true sentences respectively was confirmed by numerous subsequent studies (Gough, 1965; Slobin, 1966; Trabasso, Rollins & Shaughnessy, 1971; Clark & Chase, 1972; Carpenter & Just, 1975; Wason and Jones, 1963; Wales and Grieve, 1969, Arroyo, 1982; Vender, Delfitto, 2010). Most of these studies adopted the paradigm of the sentence-picture verification task. In this paradigm participants are presented with a sentence and subsequently with a picture. The picture presented can match (true conditions) or mismatch (false conditions) the sentence meaning, and participants are requested to evaluate if the picture correctly represents the meaning of the sentence or not, by pressing "true" or "false" key. Of course, this paradigm does not require to use sentences which are inherently true or false, as a given sentence is true or false depending on the picture it is associated with. Using this paradigm, Clark and Chase (1972, Exp.1) implemented an experiment in which they presented affirmative or negative sentences that described the disposition of a plus and of a star (e.g. "the plus is (not) above the star") and a picture that could match (true condition) or mismatch (false condition) the sentence meaning (e.g. the picture could represent 'a plus above a star' or 'a star above a plus'). In each experimental trial, a sentence and a picture were presented simultaneously on a computer screen, with the sentence presented on the left side and the picture on the right side of the screen. Participants were instructed to read the sentence first and then to compare it with the picture and to answer "true" if the picture correctly represented the sentence meaning or "false" in the other case. Clark and Chase, analyzing correct answers only, found that with affirmative sentences participants were faster to evaluate pictures that matched the sentence meaning while for negative sentences they were faster to evaluate pictures that mismatched the sentence meaning. Another famous

experiment that studied negation using a sentence-picture verification task was reported in Carpenter and Just (1975). In this experiment the authors presented sentences that described the colour of some simple geometrical shapes (e.g. circles). The sentences could be, for example, “It is true that the circles are red” or “It is not true that the circles are red”. The sentences were paired with a picture that could depict red circles or black circles. As in Clark and Chase (1982), analyzing the reaction times on correct responses in the verification task, the authors found participants were faster to answer in true rather than false conditions with affirmative sentences, but in false rather than in true conditions with negative ones. The polarity by truth-value interaction effect was also found in experiments that inverted the presentation order of the stimuli, so that the picture was presented before the sentence and participants evaluated if the sentence correctly described the depicted scene (Trabasso et al., 1971), and in experiments that asked to verify the sentences, responding “true” or “false”, against common world-knowledge, such as in “Paris is (not) a French city” vs “Paris is (not) an American city” (Wason and Jones, 1963; Wales and Grieve, 1969, Arroyo, 1982).

Electrophysiological Experiments

Above we reported some behavioural experiments that found the polarity by truth-value interaction effect. However, this effect was not only found in behavioural experiments but also using other methods. Some salient evidence of the effect came from Electroencephalography (EEG) experiments, in particular from experiments that used the Evoked Related Potentials (ERPs) technique. ERPs are changes in the EEG signal, and so in the electrical activity of the brain, due to the presentation of an event or a stimulus. ERPs are thought to reflect the cognitive processes the brain does on the stimulus presented (Zanellini, 2020). One of the most studied ERP components in language comprehension is the N400. The N400 is a negative waveform that has its pick around 400 milliseconds after the presentation of a word (Kutas & Hillyard,

1980; 1984) and is thought to reflect the difficulty of the semantic integration of a word with its prior semantic context (Luedtke, Friedrich, De Filippis, Kaup, 2008; for a review Kutas & Federmeier, 2000). In other words, the easier the integration of a word with its prior context, the smaller will be the N400 amplitude elicited by that word. The first experiment that studied negative sentences using the N400 component was reported by Fischler and colleagues (Fischler, Bloom, Childers, Roucos & Perry; 1983). In this experiment the authors presented participants with affirmative or negative sentences that might be true or false with respect to world-knowledge. The sentences were of the type:

1.a. "The robin is a bird" (True Affirmative, TA)

1.b. "The robin is a tree" (False Affirmative, FA)

2.a. "The robin is not a tree" (True Negative, TN)

2.b. "The robin is not a bird" (False Negative, FN)

The sentences were visually presented frame by frame, divided in three different frames (e.g. The robin/is (not)/a bird). The task of participants was to read the sentences and then to answer "true" or "false". The EEG signal was registered while participants read the sentences. Analyzing the reaction times on correct responses in the truth-judgment task the authors found the polarity by truth value interaction effect we presented in the above section. Indeed, they found participants were faster to evaluate affirmative true and negative false sentences with respect to affirmative false and negative true sentences, respectively. Interestingly, the ERPs results were in line with the behavioural findings. Indeed, the N400 elicited by the last frame of the sentences, which was the frame that defined the truth-value of each sentence, was lower for affirmative true and negative false sentences with respect to that elicited by affirmative false and negative true sentences, respectively. This result seems to indicate that with affirmative sentences participants were facilitated in integrating words that made the sentences true rather than false, while for negative sentences they were facilitated in integrating words

that made the sentences false rather than true. After Fischler, other EEG studies have replicated this effect of polarity by truth-value interaction on the N400 component (Luedtke et al., 2008; Scappini, Delfitto, Marzi, Vespignani & Savazzi, 2015). For example, starting from the evidence that also pictures can elicit the N400 component (Nigam, Hoffman, & Simons, 1992; Ganis, Kutas, & Serano, 1996), Luedtke et al. (2008) studied negation using the N400 and a sentence – picture verification task. In their experiment the authors presented affirmative or negative sentences of the type: "In front of the tower there is a ghost" vs "In front of the tower there is no ghost". All the sentences were visually presented word by word, with each word presented for 300ms and with an ISI between words of 300ms. Two-hundred and fifty ms after the sentence offset, the authors presented a picture that could match (true condition) or mismatch (false condition) the sentence meaning. For example, for the sentences mentioned above, the picture could represent a ghost in front of a tower or a lion in front of the tower. The participants' task was to silently read the sentences and to evaluate if the picture represented correctly or not the meaning of the sentence, by pressing one of two keys corresponding to "true" and "false". ERPs to the pictures, following affirmative sentences elicited a smaller N400 in the true conditions (i.e., matching condition) than in the false conditions (i.e., mismatching condition). For negative sentences this ERPs pattern elicited by pictures reversed: in the false conditions it elicited a smaller N400 with respect to the true conditions. The result was interpreted as evidence that with affirmative sentences participants were facilitated in interpreting pictures in the true condition, while with negative sentences they were facilitated in interpreting pictures in the false condition. This pattern is apparently coherent with the polarity by truth value interaction effect observed in the behavioural sentence-picture verification tasks reviewed above. Recently, however, Scappini and colleagues (Scappini et al., 2015) criticized the experiment by Luedtke et al. (2008), suggesting that their results could be due to simple priming effects. Indeed, Scappini and colleagues underlined as Luedtke et al.

(2008) found a smaller N400 in those conditions (affirmative true and negative false) in which the pictures represented the object mentioned in the sentences and found a larger N400 in those conditions (affirmative false and negative true) in which the pictures represented one different object to that mentioned in the sentences. To investigate if the results by Luedtke and colleagues were actually due to simple priming effects, Scappini and colleagues (2015) implemented a sentence-picture verification task experiment similar to that of Luedtke and colleagues (2008), in which they presented affirmative or negative sentences that described a Disney character as he/she was doing some actions. The sentences could be, for example, "Aladin is closing the door" vs "Aladdin is not closing the door". However, differently from the pictures adopted by Luedtke et al. (2008), in Scappini et al. (2015) 's experiment, the pictures always depicted two different characters, one mentioned in the sentence (e.g. Aladin) and one not mentioned (e.g. Jasmine). Moreover, one of the two characters depicted was doing the action mentioned in the sentence (e.g. closing a door) while the other one was doing another action (e.g. dropping a vase). In this way, for example, for the affirmative sentence: "Aladin is closing the door" the picture presented in the true condition represented Aladin closing a door and Jasmine who was dropping a vase while, in the false condition, the picture represented Aladin who was dropping a vase and Jasmine who was closing the door. For a negative sentence of the type: "Aladin is not closing the door", instead, the picture presented in the true condition represented Aladin who was dropping a vase and Jasmine who was closing a door while, in the false condition, the picture presented Aladin who was closing a door and Jasmine who was dropping a vase. It is clear how, in this experiment, in all the experimental conditions, the pictures presented always represented the character and the action mentioned, and one character and one action not mentioned in the sentences. In this way all the experimental conditions were balanced regarding the number of primed and not primed objects represented in the pictures. Anyway, analyzing the N400 results elicited by the pictures, the authors found

that their results replicated those of Luedtke and colleagues (2008). Indeed, with affirmative sentences, pictures presented in the true condition elicited a smaller N400 with respect to pictures in false conditions, while with negative sentences, pictures presented in the false conditions elicited a smaller N400 with respect to pictures in the true condition. These results seem to indicate that the polarity by truth value effect found in the experiment of Luedtke was not due to simple priming effects[1].

Explanations of the Polarity by Truth-Value Interaction Effect

We have seen that many truth-judgment experiments have found the so-called polarity by truth-value interaction effect (again, facilitation in handling true affirmatives vs. false affirmatives and false negatives vs. true negatives). Different cognitive models have been proposed to explain this result. All these models can be grouped in a class of models which has been dub "two-stage models" of negation comprehension because they all share the idea that in order to derive the actual meaning of a negative sentence (or actual state of affair) there is a first moment in which the cognitive system deals with the affirmative counterpart meaning of the sentence (or negated state of affair). Now we will see how these models explain the polarity by truth value interaction in detail.

Propositional Model

The propositional model (Clark & Chase, 1972; Carpenter & Just, 1975) is developed within a framework which assumes that the meaning of sentences is represented in a propositional

[1] Despite both actions and characters were all presented in the overall picture in affirmative true e negative false conditions both mentioned action and character were on the same side of the picture in a same scene while, for affirmative false e negative true conditions, character and action were on different sides of the screen into different scenes. Thus, perceptual advantages in picture interpretation biased upon the read sentence cannot be completely dismissed.

format which is more abstract than its linguistic representation. This model has been formulated in slightly different ways by Clark and Chase (1972) and by Carpenter and Just (1975). Both proposals aim at explaining the polarity by truth-value interaction effect in sentence-picture verification tasks. The two proposals differ for minor details but their explanation of the polarity by truth-value effect is the same; thus, for simplicity's sake, in the following we will describe in detail the Clark and Chase proposal (1972). Let's take a sentence-picture verification task experiment in which an affirmative sentence like "The star is above the plus" and a negative sentence like "The star is not above the plus" is to be compared with images of a 'a star above a plus' and 'a plus above a star' (see Figure 1) (Clark and Chase, 1972). The model of Clark and Chase suggests that both the meaning of sentences and the meaning of pictures are encoded in the same propositional format, and so a comparison between the two can be made. So, for the sentences of our example, the affirmative sentence is encoded in a single proposition such as (STAR above PLUS), while the negative one in a proposition composed by the same proposition of the affirmative, plus a false tag [false(STAR above PLUS)]. Within this proposal, the pictures are always encoded as affirmative propositions, and so, in the case of our example, they are composed by single propositions as (STAR above PLUS) and (PLUS above STAR). The model proposes that the verification task is done by comparing step by step, always starting from the inner proposition going to the most external constituent (Tian, 2014), the propositional representation of the sentence and the propositional representation of the picture. The model assumes that the verification task is more difficult more are the mismatches encountered during the comparison process of the representation of the meaning of the sentence and the picture. Thus, the model explains the polarity by truth-value interaction effect as due to the fact that there are more mismatches between the representation of the sentence and the representation of the picture in affirmative false and negative true conditions with respect to affirmative true and negative false conditions. Coming

back to our example, we have seen that both the sentence "The star is above the plus" and the sentence "The star is not above the plus", except for the "false" tag of the negative sentence, share the proposition (STAR above PLUS). This proposition of the sentences does not match with the proposition (PLUS above STAR) of the image of 'a plus above a star' which is presented in affirmative false and negative true conditions, but it matches with the proposition (STAR above PLUS) of the image of 'a star above a plus' which is presented in affirmative true and in negative false conditions (see Figure 1). In light of this, we can understand that the Clark and Chase model explains the polarity by truth-value interaction effect as due to the fact that in affirmative false and negative true conditions there is a mismatch between the proposition of the sentence representation and that of the picture that it is not present in the affirmative true and negative false conditions. For completeness' sake, it is also important to outline that the Clark and Chase model predicts that the verification task is more difficult with negative sentences with respect to affirmative ones. First, because the encoding of the negative sentence is more demanding as it requires the addition of the false tag to the proposition, and second because this false tag encoded in the representations of negative sentences is not present, and so mismatches with, the representations of the images.

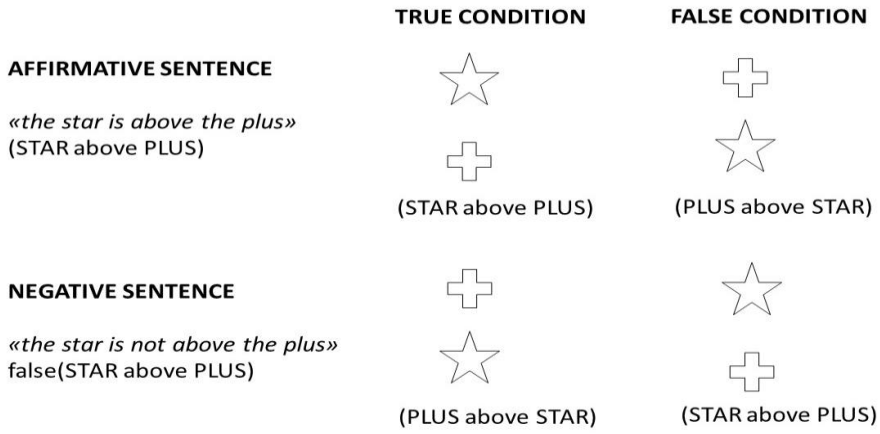


Figure 1. Graphical representation of the explanation of the polarity by truth-value interaction effect, Clark and Chase model (1972).

Two-Step Simulation Hypothesis

Another model that is suitable to explain the polarity by truth-value effect is the model proposed by Kaup and Zwann (Kaup, Yaxley, Madden, Zwaan, & Lüdtke, 2007; Zwann, 2012), known as two step simulation hypothesis (TSSH). Differently from the propositional model, the TSSH does not assume that sentences need to be represented in propositional format and that the task is performed by comparing abstract representations.

The TSSH has been developed into the framework of the embodied theory of language processing. The embodied theory assumes that in order to comprehend a sentence the system must construct a perceptual mental simulation of the sentence meaning. It is evident how one of the biggest challenges "of this view is to explain how abstract linguistic operators such as negation are represented perceptually." (Tian, 2014). So, the TSSH proposes that, while after reading an affirmative sentence the system can immediately construct an embodied mental simulation of the actual state of affairs conveyed by the sentence, the comprehension of a negative sentence requires two distinct and sequential processing steps: a first step in which the affirmative counterpart of the sentence (*the negated state of affairs*) is mentally simulated (in the format of an embodied representation) and a second step in which, the rejection of the first simulation enables to develop an embodied simulation of the actual situation described by the sentence. Leaving aside the debate about when and what is simulated after the rejection (cnfr. Zwann, 2012), it is important to underline that, according to the hypothesis, in order to interpret a negative sentence such as, for example, "the boy is not eating", the system initially must go through a mental simulation of 'a boy that is eating'. The TSSH interprets the polarity by truth-value interaction as evidence in favour of the need of developing a behaviourally costly first simulation that has then to be rejected. Indeed, in line with the propositional model (described in the previous paragraph of this thesis) but assuming the representations at play differ in nature, the TSSH assumes that the system is facilitated in the verification task when

the picture matches its (first) mental simulation. So, after reading an affirmative sentence (e.g. "the star is above the plus") the system's mental simulation will match the picture presented in the true conditions (e.g. a star above a plus) and not that presented in the false conditions (e.g. a plus above a star). On the other hand, for negative sentences, immediately after reading the sentence (e.g. "the star is not above the plus"), the (first) mental simulation will match the picture presented in the false conditions (e.g. a star above a plus) and not that presented in the true conditions (e.g. a plus above the star).

Negation and its affirmative counterpart

The substantial difference between the propositional model and the two-stage hypothesis seems to concern the nature of the sentence's mental representation (abstract vs embodied). Of course, from the evidence of the polarity by truth-value interaction effect we reported above it is impossible to define the nature of the mental representation involved in the task, and so to discern between the two proposals. More information about the nature of the mental representation involved in the processing of negation could be acquired by neuroimaging (e.g. fMRI) and neurophysiological (e.g. TMS) studies. Indeed, studies that investigate the brain areas involved during negation processing (e.g. Tettamanti et al., 2008) and studies that investigate the inhibitory mechanisms of negation (e.g. Liuzza, Candidi, Agliotti, 2011; Papeo, Hochmann, Battelli, 2016) could help in testing the two theories one against the other. Anyway, for the objective of the present thesis, it is outside of our interest to distinguish between the two proposals. The interesting point for this work is that both the propositional model and the two-step simulation hypothesis agree that the polarity by truth-value interaction effect is explainable by the fact that in the processing of negation, there is a first moment in which the negation is not considered because the system accesses the affirmative counterpart meaning of the sentence. Moreover, empirical evidence in line with this idea, and thus in support of these

models, are not only found in the explicit truth-judgment task experiments but also in many other paradigms that have been used to investigate negation comprehension. In the following section we will review some of these paradigms and their findings.

Probe-recognition task

The probe recognition task is a paradigm in which participants are asked to read some sentences that, usually, implied that a given mentioned entity (character or object) is in a specific state, configuration or position (e.g. "John put the pencil in the cup" implied that the pencil is vertically oriented (Stanfield & Zwaan, 2001)). After each sentence, a picture is presented. Based on the picture presented, the trials of the paradigm are divided in experimental and filler trials. In experimental trials, the picture presented always represents the mentioned entity (e.g. "a pencil"), while in filler trials the picture presented always represents an alternative entity (e.g. "a dog"). The participants' task is to respond if the entity represented in the picture has been mentioned in the previous sentence or not. This means that participants have always to respond "yes" in experimental trials and "no" in filler trials. However, it is important to note that in experimental trials, despite the entity is always the mentioned object, the picture can fit or not with sentence meaning in terms of specific features (pencil vertically/horizontally oriented for a "pencil in the cup"). It is interesting to note that, despite the required response in experimental trials is always "yes", reaction times show that participants are faster to respond when the position/orientation of the entity in the picture matches its position/orientation presupposed in the sentence with respect to when it does not (Zwaan, Stanfield & Yaxley, 2002; Stanfield & Zwaan, 2001; De Koning, Wassenburg, Bos, & Van der Schoot, 2017; Mannaert, Dijkstra & Zwaan, 2019). This result is interpreted as evidence of the fact that participants are faster to respond to pictures that match, rather than to pictures that mismatch, the mental representation triggered by the linguistic input. This paradigm has been clearly

developed to highlight the embodied/perceptual nature of the representation of the meaning of an utterance and can be used to check if (and when) specific features are simulated and if these representations affect simple tasks for which that feature is not relevant.

Some authors used this smart paradigm to investigate comprehension of negative sentences (Kaup, Luedtke, & Zwann, 2005; 2006; Kaup, Yaxley, Madden, Zwaan & Luedtke, 2007; Tian, Breheny and Ferguson, 2010). For example, Kaup, Luedtke and Zwann (2005) presented participants with sentences of the type "The X is above/below the Y" or "The X is not above/below the Y". They had to press a key as soon as they read the sentence. After the key pressure, a picture in which the two objects (X and Y) were one above the other was presented in two delay conditions: just after the sentence (0 ms) or after 1500 ms. In the match condition the two objects were displaced as described by the sentence, in the mismatch condition the objects were displaced in the opposite position to that described by the sentence. The authors found that for pictures presented just after the sentences, in the case of affirmative sentences, participants were faster to respond to the pictures in which the objects mentioned were displaced as described by the sentence (match condition), whereas in the case of negative sentences they were faster to answer in the case of pictures in which the objects were displaced in the opposite state to that described (mismatch condition). For pictures presented after a 1500 ms delay, the authors found the same pattern in the case of affirmative sentences, that is participants were still faster to respond in the match condition while, in the case of negative sentences, participants were equally fast to respond to pictures in which the objects were displaced as described or in the opposite position to that described by the sentence. The result with negative sentences when the pictures were presented 0ms after the sentences, indicates that, in this delay condition, participants were facilitated in responding to pictures that represented the objects displaced as described by the affirmative counterpart meaning of the sentences (a X above a Y for the sentence "the X is not above the Y") rather than to pictures

that represented the objects displaced as described by the actual meaning of the sentences (a Y above a X for the sentence "the X is not above the Y"). This result was interpreted by the authors as evidence that, initially, the mental representation triggered by the negative sentences corresponded to that elicited by the affirmative counterpart meaning of the sentence. So, the authors concluded that, in the interpretation of the negative sentences, participants initially simulated the affirmative counterpart meaning of the sentences. The result of faster answers in the mismatch vs the match condition with negative sentences, was subsequently replicated by other experiments that used the probe-recognition task. For example, Kaup and colleagues (Kaup, Yaxley, Madden, Zwaan, & Luedtke, 2007) implemented an experiment in which they presented negative sentences such as, for example, "There was no eagle in the sky". Two hundred-and-fifty ms after the sentence, a picture was presented in which the object mentioned (e.g. eagle) was depicted in a position coherent to that implied by the sentence (e.g. an eagle with nested wings – match condition) or in a position incoherent to that implied by the sentence (an eagle with stretched wings in flying position – mismatch condition). Analyzing the reaction times, the authors found that participants were faster in responding to pictures in which the object was depicted in the incoherent position to that implied by the sentence vs pictures in which the object was depicted in a coherent position to that implied by the sentence. Similarly, Tian, Breheny and Ferguson (Tian, Breheny & Ferguson; 2010) implemented an experiment in which presented negative sentences that described an object in a physical state that can have an opposite (e.g. "Joe didn't close the door"). Two hundred-and-fifty ms after the sentence, they presented a picture in which the object mentioned (e.g. the door) was depicted in the physical state described by the sentence (e.g. open – match condition) or in the opposite physical state (e.g. closed – mismatch condition). Analyzing the reaction times, the authors found participants were faster in responding to pictures that depicted the object in the opposite physical state to that described (mismatch condition). As the results by Kaup et al., (2005), also

the results by Kaup et al. (2007) and by Tian et al. (2010), indicate that with negative sentences participants were facilitated in responding to pictures that depicted the object in the position implied by the affirmative counterpart meaning of the negative sentences and so they can be interpreted as evidence of the occurred access to affirmative counterpart meaning of the sentence in the interpretation of the negation. One could speculate that these results found with negative sentences within the probe recognition paradigm were simply due to the fact participants deliberately ignored the presence of the negation in the sentences. However, we think it is important to outline that all these experiments included a task to discourage participants from doing this. For example, one of the tasks often used was that to randomly present after some trials a question about the last sentence read to which, in order to answer correctly, a fully comprehension of the sentence was required. This means that to answer the question correctly participants had to take negation into account if there was any. Usually, participants answered these questions with an accuracy clearly above chance, showing that they did not deliberately ignore the presence of the negation in the sentences.

Lexical Decision Task

In lexical decision tasks participants are usually presented with a series of individual words and non-words and they are asked to judge whether each item is a word or not, responding "yes" or "no". Interestingly, it has been found that participants are faster to categorize a word as such if the previous trial involved another word that is semantically associated with it (Fischler, 1977a, Fischler, 1977b, Blank, 1980, Perea & Rosa, 2002). In other words, it has been found that participants are faster to judge that "bark" is a word if the previous trial involved the word "dog" rather than the word "sky". This result is thought to be due to the fact that processing a word automatically activates the concepts that are semantically associated with it, increasing the availability of the associated concepts in the semantic memory. This

effect is called semantic priming and it has been found both in tasks which explicitly require access to word meaning (e.g. semantic categorization) and also in tasks such as lexical decision which do not explicitly require the retrieval of the meaning of a word to be accurately performed. Some authors adapted the lexical decision task to investigate how negative sentences are comprehended (Giora, Balaban, Fein & Alkabets, 2004; Hasson & Glucksberg, 2006). For example, Giora et al., (2004) presented participants with sentences of the type:

3 The instrument is sharp.

4 The instrument is not sharp.

One hundred ms after the sentence the authors presented a stimulus that might be a word or a non-word. The task of participants was to categorize the stimulus presented as a word or a non-word, by responding "yes" or "not". It is important to outline that half of the words presented were strongly related to the last word presented in the sentences, while the other half were unrelated to the last word. That is, for the example sentences above, the related word presented was "piercing" while the unrelated word was "leaving". The authors found that participants were faster to recognize the word as such when it was related to the last word of the sentence (e.g. "piercing") with respect to when it was unrelated (e.g. "leaving"), independently from whether the sentence was presented in the affirmative or in the negative form. This result seems to suggest that the negative sentence activated the same concept activated by the affirmative sentence, and so that participants, in the moment to perform the task, did not take the presence of the negation in the sentence into account. The result was interpreted by the authors as evidence that immediately after a negative sentence, participants accessed its affirmative counterpart meaning. Hasson and Glucksberg (Hasson and Glucksberg, 2006), however, criticized the results by Giora and colleagues (2004) because in their view the stimuli used did not allow to disentangle if the effects found were due to the fact that participants accessed the affirmative counterpart meaning of the negative sentences or to lexical priming effects caused

by the strong relation between the mentioned and the target word in one case (sharp - piercing), and the weak relation between the mentioned and the target word in the other case (sharp - leaving). Thus, they decided to run an experiment using the same paradigm adopted by Giora et al. (2004) but presenting metaphors instead of simple sentences. Indeed, authors proposed that metaphors have the advantage that it is possible to present target words that are primed by the meaning of the metaphors but that are not primed by the lexical items that constitute the metaphors. For example, they propose that the metaphor "*This lawyer is a shark*" primes the word *vicious*, but that the word *shark* alone does not (Blasko and Connine, 1993, exp. 5; Glucksberg et al., 2001). In their experiment the authors presented some affirmative or negative metaphors of the type:

5 Some surgeons are butchers.

6 Some surgeons are not butchers.

After these metaphors, the authors presented a word or a non-word and participants had to respond if the item presented was a word or not. Half of the words presented were words related to the meaning of the affirmative metaphors, while the other half were words related to the meaning of the negative metaphors. Thus, for the two examples above, the word related to the affirmative metaphor was *clumsy*, while the word related to the negative metaphor was *precise*. These targets words were presented in three delay conditions: 150ms, 500ms or 1s after the metaphor. Analyzing the reaction times in the task the authors found that in the 150ms and 500ms delay conditions participants were faster to recognize as words the words that were related to the affirmative metaphor, for both affirmative and negative metaphors. In the 1s delay condition, instead, participants were faster to respond to words related to the affirmative metaphor after the affirmative metaphor, but they were equally fast to respond to words related to the affirmative or to the negative metaphor, after the negative metaphor. Based on the evidence that in the 150ms and 500ms delay conditions affirmative and negative metaphors

seemed to activate the same concepts, the authors concluded that the interpretation of a negative sentence requires a first moment in which the affirmative counterpart meaning of the sentence is accessed, thus confirming the main conclusion of Giora et al. (2004) with a more apt paradigm which exclude explanations in terms of bare activation of the meaning of single words, requiring to assume that utterance meaning drives such effects. As well as for the probe-recognition paradigm, also in these lexical decision tasks one can speculate that the results with negative sentences can be explained by the fact that participants deliberately choose to ignore the presence of the negation in the sentences. However, both Giora et al., (2004) and Hasson & Glucksberg (2006), in their experiments randomly presented some comprehension questions to which, in order to answer correctly, a full comprehension of the sentences presented was required. The fact that participants answered these questions with an accuracy clearly above chance shows that they did not ignore the negation; moreover, data from the experiment with 1s delay by Hasson & Glucksberg (2006) further confirm that negation was not ignored.

Visual World Paradigm

Other interesting evidence in line with the idea that the interpretation of negation requires a moment in which the affirmative counterpart meaning of the sentence is activated, came from the eye-tracking method and the visual world paradigm (Orenes, Beltrán, & Santamaria, 2014, Nordmeyer & Frank, 2014). In this paradigm participants are presented with a visual workspace in which some objects are depicted. During the presentation of the visual workspace participants are asked to listen to a sentence about the content of the workspace, and their eye movements are recorded. It has been found that when linguistic input and visual input match, the eyes automatically move towards the corresponding visual input on the workspace (Cooper, 1974; Tanenhaus, Spivey-Knowlton, Eberhard & Sedivy, 1995). Thus, eye movements are thought to reflect the mental representation triggered by the linguistic input. Some authors used

this paradigm to investigate the comprehension of negative sentences. For example, Orenes and colleagues (2014) presented participants with a visual workspace composed by four figures of different colours (red, green, yellow, blue). During the presentation of this workspace, they asked participants to listen to sentences of the type: "The figure is red" or "The figure is not red". Importantly, before these sentences, participants heard another sentence such as: "the figure could be red or green or blue or yellow" (quaternary condition), or "the figure could be red or green" (binary condition), thus restricting the pool of the figures to which pay attention to only two. The task of participants was simply to listen to the sentences and look at the visual workspace. Analyzing the looks of participants after the critical sentences (e.g. "The figure is (not) red") the authors found that, immediately after the sentences, participants moved their eyes toward the named colour (e.g. red), irrespective of whether the sentence presented was affirmative or negative. However, in the binary condition, when the critical negative sentences were preceded by a sentence that restricted the salient figures to only two, the number of fixations toward the named colour did not result significantly higher than the number of fixations toward the figure with the other colour to that mentioned and, after around 1300ms, participants redirected their looks toward the figure with the other colour. In the quaternary condition, when the negative sentences were preceded by a phrase that did not restrict the pool of salient figures, instead, the number of looks toward the mentioned colour resulted immediately significantly higher than the number of looks toward the non-mentioned colours, and participants kept their gaze on the mentioned colour without redirecting their gaze to the other three non-mentioned figures. In light of the fact that in the binary condition there were not statistically significant higher fixations to the mentioned vs the other colour and that around 1300ms participants redirected their looks toward the non-mentioned colour, the authors interpreted their results as against the idea that, in this condition, the affirmative counterpart meaning of the sentence was accessed while, the results in the quaternary condition, were

interpreted as in line with this idea. We think, however, that also the result in the binary condition cannot be taken as against the idea of the access to the affirmative counterpart meaning of the sentence. Indeed, also in this case, despite failing to reach the statistical significance, initially participants moved their eyes toward the named colour and only after 1300 ms they were able to redirect their gaze toward the other colour. So, we think that, overall, the results by Orenes and colleagues (2014) can be interpreted as evidence in support of the idea that in the interpretation of negative sentences, at least initially, the affirmative counterpart meaning of the sentence is accessed. Additional evidence of the initial access to the affirmative counterpart meaning of the sentence during negation interpretation can be found in an experiment by Nordmeyer et al. (2014, Experiment 2). In this experiment the authors presented sentences of the type "Look at the boy with apples" or "Look at the boy whit no apples", and the visual workspace was composed by a boy with two apples on one side of the screen and a boy with two other items (e.g. a boy with two boxes) on the other side of the screen. Analyzing the look of participants, the authors found that, after the sentences, participants moved their eyes toward the boy with the items mentioned in the sentence, when the sentence was affirmative as well as when it was negative.

Additional Evidence

Additional evidence in line with the idea that the interpretation of a negative sentence requires a stage in which the affirmative counterpart meaning of the sentence is accessed come from an experiment by Mayo and colleagues (Mayo, Schul, & Burnstein, 2004) and from an experiment by Dudschig and colleagues (Dudschig, Mackenzie, Maienborn, Kaup & Leuthold, 2019). Mayo et al. (Experiment 1) implemented a paradigm that probably, among the experimental paradigms we have seen until now, is the most similar to a truth-judgment task. In their experiment the authors presented a sentence that described a person. This sentence could be

presented in its affirmative version (e.g. “Tom is a tidy person”) or in its negative version (e.g. "Tom is not a tidy person"). After the sentence, the authors presented another sentence that described a behaviour. This second sentence could describe a behaviour congruent or incongruent with the description of the person made in the first sentence. Of course, a behaviour description (e.g. “Tom’s clothes are folded neatly in his closet”) that was congruent with the description of the person in the affirmative version (e.g. “Tom is a tidy person”), was incongruent with the description of the person in the negative version (e.g. "Tom is not a tidy person"). While, at the opposite, a behaviour description (e.g. “Tom forgets were he left his car keys”) that was incongruent with the description of the person in the affirmative version (e.g. “Tom is a tidy person”), was congruent with the description of the person in the negative version (e.g. "Tom is not a tidy person"). The task of participants was to evaluate if the behaviour description was congruent or incongruent with the description of the person. The authors found that if with the affirmative version of the person’s description participants were faster and more accurate in evaluating congruent vs incongruent behaviour descriptions, with the negative version they were faster and more accurate in evaluating incongruent vs congruent behaviour descriptions. This result indicates that participants were always faster to respond to the behaviour description that was congruent with the affirmative description of the person, independently of the version (affirmative vs. negative) of the description sentence. The result was interpreted by the authors as in line with the proposal of the two stage models of negation comprehension, and so as evidence of the occurred access to the affirmative counterpart meaning of the sentence during the interpretation of the negation. The other experiment we think is worth to mention is an EEG experiment by Dudschig et al. (2019). In this experiment the authors analysed the N400 component to investigate the processing of some affirmative sentences and of their respective negative forms. The authors presented participants with sentences that in the affirmative form could be correct (e.g. “Zebras are (not) stripy”), world-

knowledge violations (e.g. “Ladybirds are (not) stripy”) or selectional restriction semantic violations (e.g. “Thoughts are (not) stripy”). The task of participants was simply to listen to the sentences and to answer some comprehension questions about the sentences they were listening to. The authors found that the N400 elicited by the last word of the sentences was larger for the world-knowledge violation and semantic violation sentences with respect to that elicited by the correct sentences. Moreover, the N400 results were not influenced by the fact that the sentences were presented in the affirmative or in the negative version. This result seems to suggest that the presence of the negation did not affect the processing of the sentences at the level of the N400 component. The authors interpreted the result as evidence that in the processing of the negation participants did not take immediately the negation into account and so that they initially accessed the affirmative counterpart meaning of the sentence.

Negation meaning is immediately accessible

Until now we have reviewed pieces of evidence in line with the idea put forward by the two stage models of negation comprehension, that, in order to interpret a negative sentence, the system must pass through a first moment in which the meaning of the affirmative counterpart of the sentence is accessed. However, other authors (Nieuwland & Kuperberg, 2008; Dale & Duran, 2011), challenged this view because in their experiments they found evidence that suggest that, at least in specific situations, it is possible to access immediately the actual meaning of a negation, without necessarily pass through the first stage proposed by the two-stage models of negation comprehension. In this section we will review this evidence.

The Pragmatic View

The most influential experiment that has challenged the idea that the interpretation of the negation requires a moment in which the affirmative counterpart meaning of the sentence is

accessed is an EEG experiment by Nieuwland & Kuperberg (2008). The authors, starting from the evidence that in everyday communication negation is understood without any problem, proposed that to be comprehended easily negative sentences require a pragmatic context in which the use of a negation makes sense. Indeed, they proposed that, differently from affirmative sentences, out of a pragmatic context, negative sentences result under-informative and so not felicitous. On the basis of this assumption, they highlighted as the experiments that found evidence in line with the idea of the access to the affirmative counterpart meaning of the negation presented negative sentences out of a pragmatic context, in isolation, and thus sentences that were not felicitous. To test if their intuition about the importance of the pragmatic context in the interpretation of the negation was correct, in their EEG experiment, the authors compared the processing of negative sentences that could be uttered in a pragmatic context that made them felicitous, or in a pragmatic context that made them not felicitous. To do so they constructed some sentences in the affirmative and in the negative version that contained a bipolar predicate (e.g. safe/dangerous). They asked to an independent group of participants to judge if the sentences they constructed sounded natural within a conversation and to evaluate if the sentences were true or false. On the basis on the naturalness ratings with the negative version of the sentences the authors divided the sentences in pragmatically licensed/felicitous and pragmatically unlicensed/not felicitous. The sentences were also divided, on the basis of the truth-value judgment, in true or false. After the pre-tests the authors obtained sentences of the following types:

Pragmatically licensed (felicitous):

7.a. With proper equipment, scuba diving is really <u>safe</u> and often good fun (True Affirmative)

7.b. With proper equipment, scuba diving is really dangerous and often good fun (False Affirmative)

8.a. With proper equipment, scuba diving is not really dangerous and often good fun (True Negative)

8.b. With proper equipment, scuba diving is not really safe and often good fun (False Negative)

Pragmatically unlicensed (not felicitous):

9.a. Bulletproof vests are very safe and used worldwide for security (True Affirmative)

9.b. Bulletproof vests are very dangerous and used worldwide for security (False Affirmative)

10.a. Bulletproof vests are not very dangerous and used worldwide for security (True Negative)

10.b. Bulletproof vests are not very safe and used worldwide for security (False Negative)

In the experiment, the sentences were visually presented word by word in the centre of a computer screen, and the task of participants was to silently read the sentences. In order to ensure participants paid attention to the sentences content, after a variable number of sentences a word was presented and the task of participants was to evaluate if that word was semantically related or not to the sentence just read. For affirmative sentences, the N400 elicited by the critical word of the sentences, and so the bipolar adjective that made the sentence true or false (e.g. safe/dangerous), was smaller for affirmative true sentences with respect to affirmative false sentences. For negative sentences, instead, in the pragmatically licensed condition true sentences elicited a smaller N400 with respect to false sentences while in the pragmatically

unlicensed condition true sentences elicited the same N400 as false sentences. The results with negative sentences in the pragmatically licensed condition show that, differently from the other experiments that investigated the N400 component in the processing of negative sentences (e.g. Fischler et al., 1983), participants were facilitated in integrating the word that made the sentence true with respect to false. This result was interpreted by the authors as against the proposal of the two-stage models of negation comprehension which foresee a facilitation for false negative sentences with respect to true ones, and so as evidence that, in the pragmatically licensed condition, participants were able to access immediately the actual meaning of the negation. The authors concluded that to the extent that a negative sentence is interpreted within a pragmatic context that makes it felicitous, negation is not different from affirmation, and it is possible to access immediately its actual meaning. To a similar conclusion of that of Nieuwland & Kuperberg (2008) arrived also Dale & Duran (2011). In their paper Dale & Duran (2011) implemented three experiments using the sentence verification task and the mouse tracking methodology. In the experiments the authors presented affirmative and negative sentences that could be true or false with respect to world knowledge, as for example:

11.a. "Elephants are large" (True Affirmative, TA)

11.b. "Elephants are small" (False Affirmative, FA)

12.a. "Elephants are not small" (True Negative, TN)

12.b. "Elephants are not large" (False Negative, FN)

The sentences were visually presented at the bottom of a computer monitor word by word. For passing from a word to the subsequent one, participants had to click with the mouse on a button positioned at the bottom centre of the monitor. After reading the sentence participants had to judge if it was true or false, using the mouse of the computer to reach, from the bottom of the monitor, one of the two buttons labelled as "TRUE" or "FALSE" presented on the top of the monitor, one on the left and one on the right. The trajectories participants made to reach the

button were registered. The idea behind the study of mouse trajectories is that these ones capture the cognitive dynamics participants made during formulation and execution of a response (Fischer & Hartmann, 2014; Freeman & Ambady, 2009; Song & Nakayama, 2009; Spivey & Dale, 2004). Dale and Duran proposed that if a negative sentence is interpreted as proposed by the two stage models of negation comprehension, they should have found a flip in the mouse trajectories in the moment at which from the affirmative counterpart meaning of the sentence participants accessed the actual meaning of the sentence. Indeed, the idea of the authors, following the logic of the two stage models, was that if initially negation is not taken into account, a negative false sentence (e.g. "Elephants are not large") should be interpreted as a true sentence, while a negative true one (e.g. "Elephants are not small") as a false one. So, the authors predicted that, if negative sentences are interpreted in two stages, they would have found more deviations and acceleration in the mouse trajectories with negative sentences with respect to affirmative ones, because of the first moment in which participants would have moved the mouse toward the wrong response button for the truth-value judgement.

In the first experiment sentences as in (11, 12) where presented in isolation, out of a pragmatically supportive context. Mouse trajectories showed that with negative sentences participants made more deviations and accelerations/decelerations than with affirmative sentences. Moreover, though with affirmative sentences participants did not show a difference between true and false sentences, with negative sentences the authors found an increasing of deviations and acceleration/deceleration in judging true sentences with respect to false ones. The authors interpreted the results as in line with the proposal of the two-stage models of negation comprehension, and so as evidence that in the interpretation of the negation participants passed through a first moment in which the affirmative counterpart meaning of the sentence was accessed. In the second experiment the authors always presented the sentences of the first experiment but before of them they presented an introductory context that they thought

would have made the negative sentences felicitous. For example, for the sentence "Elephants are not small" the introductory context was a question such as "You want to lift an elephant?". Their prediction was that this time, in line with the proposal of Nieuwland and Kuperberg (2008), participants would have accessed immediately the actual meaning of the negative sentences and so that the mouse trajectories of the negation would have been as those of the affirmative sentences. However, despite the introductory context, analyzing the mouse trajectories, the authors found that the results replicated those of the first experiment. This result seems to be in contrast with the results by Nieuwland and Kuperberg (2008) because despite the introductory context that the authors thought would have made the negative sentences felicitous, the negative sentences were interpreted as when they were presented in isolation. However, Dale and Duran proposed that a significant difference between their experiment and that of Nieuwland and Kuperberg (2008) was that their sentences were generally shorter than those used by them. So, in the third experiment, they decided to embed the sentences in a longer introductory context such as, for example, 'You want to lift an elephant?' the mother said to her child, 'but elephants are not small'. In this third experiment the authors found that as for affirmative sentences, also for negative sentences, there were not differences in the mouse trajectories between true and false sentences. Moreover, with negative sentences participants did not make more deviations and accelerations/decelerations than with affirmative sentences. This result was interpreted by the authors as evidence in line with the proposal by Nieuwland and Kuperberg (2008), and so as a proof that when interpreted within a supportive enough pragmatic context, negative sentences are not different from affirmative sentences and so their actual meaning can be immediately accessed.

Conclusion

In this introductory chapter we have seen many pieces of evidence that seem to suggest that negation is more difficult to handle with respect to affirmation. Moreover, we have reviewed some empirical results, such as for example the polarity by truth-value interaction effect found in truth-judgment tasks, that seem to be in line with the proposal of the two stage models of negation comprehension. According to these models, during the interpretation of a negative sentence, before to access the actual meaning of the sentence, there is a first moment in which the affirmative counterpart meaning of the sentence is accessed. We have seen that these models differ each other on the assumptions regarding the nature of the mental representation involved in the interpretation of negative sentences (propositional vs embodied) but, as written above, distinguishing between the two types of models is not the topic of this thesis. More interestingly for our work is that there is empirical evidence suggesting that in some circumstances it is possible to access immediately the actual meaning of a negation, without necessarily access the affirmative counterpart meaning of the sentence (Nieuwland & Kuperberg, 2008, and Dale & Duran, 2011).

Chapter 2: Aim of the Thesis

Pragmatic Felicity

The experiments by Nieuwland & Kuperberg (2008) and Dale & Duran (2011), clearly show that there are situations in which it is possible to access immediately the actual meaning of the negation. In particular the authors proposed that this is possible when negative sentences are uttered and interpreted within pragmatic contexts that made negative sentences felicitous. However, it is not really clear which were the pragmatic factors that defined the felicity of the negative sentences in those experiments. Indeed, as we have seen, Nieuwland and Kuperberg (2008) judged a negative sentence as felicitous or not just operationally, asking to a group of participants to evaluate the naturalness of the sentences within a conversation. On the other hand, Dale and Duran (2011) seem to suggest that the length of the context may be the determinant factor that enables to access or not the actual meaning of the negation, though in their paper they do not emphasize or discuss theoretically this point. We think that, in order to bring further evidence in support of the idea that pragmatic factors can allow to access immediately the actual meaning of the negation, it is important to detail accurately the conditions in which the meaning of a negative sentence can be immediately accessed.

The investigation of this point is the main objective of this thesis. It is important to outline, however, that a step toward this direction has been already made by Tian, Breheny and Ferguson (2010, 2016) and an interesting result in line with our objective was found by Arroyo (1982). The works of these authors will be subsequently presented in this thesis and their proposals will be empirically tested.

Predictability Effects vs Pragmatic Effects

Before to present our experimental works, however, we want to point out that in our investigation we will try to avoid that possible evidence in support of the idea of an immediate interpretation of the negation for a given pragmatic factor could be attributed only to lexical predictability effects. Indeed, there is evidence in the literature suggesting that lexical predictability effects, rather than pragmatic effects, can account for some results interpretable as in line with the idea of an immediate access to the actual meaning of the negation. For example, Luedtke & Kaup (Lüdtke & Kaup, 2006, Experiment 2), compared the reading times of negative sentences with the reading times of their affirmative counterparts when the sentences were uttered in a context that weakly vs. strongly implied the likely negated word presented in a sentence. The authors found that in the context in which the negated word (e.g., *dirty*) was weakly implied, participants read negative sentences (e.g. "The t-shirt is not dirty") slower than their affirmative counterpart (e.g. "The t-shirt is dirty") while, in the context in which the negated word (e.g. *dirty*) was strongly implied, participants read negative sentences (e.g. "The t-shirt is not dirty") as fast as their affirmative counterpart (e.g. "The t-shirt is dirty"). The result in the weakly implied condition can be interpreted as in line with the idea of the two-stage models and so as evidence of the extra processing negative sentences require with respect to affirmative ones. This view is, however, difficult to reconcile with the results in the strongly implied condition since in the latter condition there is no evidence of a supposed extra processing required by negative sentences. Thus, the results in the strongly implied condition suggest that in this case participants immediately accessed the actual meaning of the negation, but it may also be compatible with an alternative explanation assuming that the two stages are performed at an anticipatory level. Another result that seems to indicate that if it is somewhat possible to predict the negated word the actual meaning of the negation can be immediately accessed is by Glenberg et al. (1999). As in Luedtke and Kaup (2006), Glenberg and colleagues

(1999) studied the reading times of negative sentences. In their experiment the authors compared the reading times of negative sentences with those of their affirmative counterparts in a context that anticipated the content of the negated word or in a context that did not anticipate it. For example, a sentence such as "The couch is not black" was presented after the context "She wasn't sure if a darkly colored couch would look the best or a lighter color" (anticipating context) or after the context "She wasn't sure what kind of material she wanted the couch to be made of" (not anticipating context). The authors found that in the context that anticipated the negated word negative sentences were read as fast as affirmative ones while in the context in which the negated word was not anticipated participants read negative sentences slower than affirmative ones. As the results by Luedtke & Kaup (2006), also the results by Glenberg et al. (1999) seem to suggest that when it is somewhat possible to predict the negated word it is also possible to access immediately (or proactively) the actual meaning of the negation. However, the effect is not always consistent. Indeed, it is interestingly to note that in another experiment, Luedtke & Kaup (2006, Experiment 1), failed to see the elimination of the costs in terms of reading times for negative sentences with respect to affirmative ones when the negated word was predictable. In that experiment the authors compared the reading times of negative sentences of the type "The water was not warm" with those of their respective affirmative counterparts in a context in which the negated word ("warm") was previously introduced by the context (e.g "Danielle wondered whether the water would be warm.") or not (e.g. "Danielle wondered what the water would be like"). Despite the reduction of the difference in terms of reading times between negative and affirmative sentences in the context in which the negated word was introduced, negative sentences were always read slower than affirmative sentences. Of course, this result is at odds with the results of the other two experiments and questions the reliability of the effects of the predictability of the negated word on the

interpretation of the negation. Thus, as stated above, we will try to control for predictability in order to avoid possible confounds.

In the same vein, we will try to control or explicitly manipulate for the possibility to predict the affirmative meaning of a negative sentence, for example the possibility to predict that "not green" means "orange". Indeed, despite a number of findings seem to suggest this is not the case (Darley, Kent & Kazanina, 2020; Tian et al, 2010; Dale & Duran, 2011, Calignano, 2020), some authors suggest that when it is possible to predict the affirmative meaning of a negative sentence it is possible to access the actual meaning of the negation without passing through the affirmative counterpart meaning of the sentence (Mayo et al., 2004; Orenes et al, 2014; Trabasso et al., 1971; Anderson, Huette, Matlock & Spivey, 2010).

Conclusion

In this chapter we presented our main Research Question which is to detail the way pragmatic felicity of an utterance can allow an immediate interpretation of negative sentences and we outlined that in doing this we will try to particularly control for predictability in order to avoid potential confounds. In most of the studies we will employ simple abstract objects like coloured shapes; in doing this we will clearly not be able to discriminate between propositional and embodied/perceptual versions of the two-stage models which has been largely debated in the last years through neuroimaging methods. For most of the experiments run here the issue at stake is if the negation can be interpreted immediately or if it needs to go through two stages, independently from the fact that the involved representations are formal propositional (Clarke) or embodied simulations (Kaup). In the next four chapters we will present the experiments we did and the factors we tested, all proposed in the linguistic and psychological literature as factors that should make a negative sentence felicitous. The first factor is the use of the negation

to refer about a minority; the second factor is the use of the negation to describe something that non-exists; the third and the fourth factors concern, respectively, the Question Under Discussion and the fact that the negation is used to describe something exceptional.

Chapter 3: Minority[2]

Introduction

In the first empirical experiment of this thesis, we decided to study whether referring about a minority is one of those pragmatic factors that has been proposed to allow an immediate access to the actual meaning of a negative sentence. In the literature, indeed, two factors that are frequently reported as factors that should make a negative sentence felicitous are the fact that the sentence is used to talk about deviations from an expectancy or the fact that the sentence is used to talk about exceptions (Wason, 1965; Givon, 1978; Colston, 1999). One of the most cited works in support of this idea, is the paper by Wason (1965). In this paper the author developed, and empirically tested, two hypotheses about two situations that make plausible to use a negative sentence. The first hypothesis, known as *exceptionality hypothesis*, proposes that *"Given a set of similar stimuli, X1, X2, ...,Xn, and a stimulus, Y, which is perceived to differ from these in one important attribute, it is more plausible to assert that Y is not X than to assert that Xi is not Y"* (Wason, 1965, pg. 8). To test this hypothesis Wason ran an experiment in which he presented participants with a picture of eight circles for three seconds. These circles were numbered from one to eight and coloured differently: seven circles with the same colour (e.g. blue) and one circle of a different colour (e.g. red). After the presentation of the picture, Wason asked participants to describe the picture aloud, specifying the number of the circle with the different colour, as for example "The circle number 4 is red, and the rest are blue". Following this description participants had to complete as faster as possible a sentence fragment that could be either affirmative (as in 1) or negative (as in 2), pronouncing in a microphone just the colour word (blue or red) that was required to make the sentence true.

[2] This study has been carried out in collaboration with Chiara Finocchiaro, Veronica Mazza and Francesco Vespignani

1.a. “The circle number 4 is …”

1.b. “The circle number 7 is …”

2.a. “The circle number 4 is not …”

2.b. “The circle number 7 is not …”.

Wason (1965) found that participants were faster to complete the negative version of the sentence (2) when this was referred to the circle with a different colour with respect to when it was referred to one of the circles with the same colour. This result is in line with the *exceptionality hypothesis* because it indicates that participants were facilitated in completing the negative sentence when it referred to the odd stimulus. The second hypothesis proposed by Wason is the *ratio hypothesis*, which can be formulated as follows: *“Given two sets of stimuli which differ considerably in magnitude, it is more plausible to deny that the smaller set possesses a property of the larger set than to deny the converse”* (Wason, 1965, pg. 8). To verify this hypothesis the author ran an experiment similar to the previous one. Indeed, he presented for 3 seconds exactly the same picture as in the first experiment described above (7 circles of one colour and one of a different colour), with the only difference that the circles were not numbered. After the presentation of the picture, Wason asked participants to describe the picture aloud, specifying the numerosity of the circles for each colour, as for example “one circle is red and seven are blue”. Finally, he presented a sentence, that might be affirmative (as in 3) or negative (as in 4), that participants had to complete as faster as possible.

3.a. “Exactly one circle is…”

3.b. “Exactly seven circles are …”

4.a. “Exactly one circle is not …”

4.b. “Exactly seven circles are not …”

Wason found that participants were equally fast to complete the negative sentences (4) when these were referred to the circle with the different colour or to the seven circles with the same colour. On the other hand, the *ratio hypothesis* would have predicted an advantage for the minority case, as in the previous experiment. While in the first experiment negative sentences resulted facilitated when the scope of negation is an individual (numbered) object, the second experiment shows that a similar effect does not clearly emerge (by correctly inferencing upon null results) in dealing with negative sentences which referred to a smaller set. So, from a shallow reading of the paper one could conclude that the *exceptionality hypothesis* was confirmed but not the *ratio hypothesis,* and so that a negative sentence is plausible if used to talk about one “odd” item but not if it used to talk about a minority of items. Wason himself, discusses the results in a rather more complex way, stating that a null result cannot be taken as a falsification of the *ratio hypothesis,* which despite not receiving a positive experimental confirmation may still hold in terms of a linguistic hypothesis of negation licensing.

Since the picture was the same in the two experiments except for numbering, Wason suggests that the more relevant difference between the two experiments, which can justify the contrasting results, is in the way participants described the picture. In details, he proposes that, in the first experiment, by describing the picture as “The circle number 4 is red and the rests are blue”, participants coded the picture as composed by one unique set of elements in which one of them is different; while, in the second experiment, describing the picture as “one circle is red and seven are blue”, participants coded the picture as composed by two separate and distinct sets, one of seven elements and one of one element. Following this line of reasoning, the author put forward the idea that the paradigm implementation is not correctly testing the ratio hypothesis since if the two sets are conceptualized as separate sets, each of them, despite

the different numerosity (1 and 7), has all the objects of the same colour and the ratio would emerge only if these were considered as a unique set.

We fundamentally agree with Wason and we think that in order to set-up the crucial manipulation of the experiment, making the numerosity comparison between the two sets conceptually relevant, the stage of the task implemented by Wason by requiring participants to exactly describe the number of elements for each set can be counterproductive, aiding the conceptualization of two different and separate sets which may not be to be compared. We thus think that a more suitable method to implement a numerosity comparison, can be that of presenting the two sets visually and simultaneously without asking any verbal description of them, as it is done in experiments that investigate the perception of magnitude (e.g. Tokita, Ishiguchi., 2013). Given the reasons just explained, it seems worthful to keep considering the hypothesis that negation is pragmatically licensed also if used to refer about a minority and not only to an exception and we thus decided to try to develop a different paradigm to test if referring to a minority allows to access immediately the actual meaning of a negative sentence. In order to do that, we implemented a sentence-picture verification task. In our paradigm participants read affirmative and negative sentences and must evaluate if a subsequent picture correctly represents (true condition) or not (false condition) the sentence meaning. The sentences talk about the colour of a group of shapes that in the picture can be represented as the majority or the minority of the shapes depicted. The majority group has twenty-four elements while the minority one six elements. In the picture, the elements of the two groups of shapes are presented intermixed and in random positions. We think that the idea to use numerosity above the subitizing range, that make difficult to exactly count the elements of each group if not required, and the intermixed and random presentation of the shapes can be a good way to overcome the limitation of the Wason paradigm, and so to lead participants to code the two group of shapes as one unique set. Having said that, with respect to negative sentences, we

have two main predictions. First of all, we predict that, if the *ratio hypothesis* is true, we will find a reduction of the costs for negative true sentences when the sentences refer to the minority, with respect to when they refer to the majority. Of course, we cannot exclude that an effect of this type could be simply due to a general facilitation in performing the task when sentences refer to the minority so, in order to exclude this possibility, we will control for numerosity effect on affirmative true sentences. This means that we will consider a result in support of the *ratio hypothesis* only in the case in which the facilitation for true sentences when they refer to the minority vs the majority is limited to negative sentences, and thus is not found with affirmative true sentences. On the basis of the *ratio hypothesis*, we additionally predict that if there is an immediate access to the actual meaning of the negation when it refers to the minority and a two-stage interpretation of the negation in the case in which the negative sentence refers to the majority, participants will be faster and/or more accurate in responding in true conditions with respect to false ones when the negative sentence refers to the minority but they will be faster and/or more accurate in responding in false conditions with respect to true ones when the negative sentence refers to the majority. As we have seen in the introduction chapter of this thesis, indeed, in truth-judgement tasks a result of faster and/or more accurate responses in false vs true conditions with negative sentences is interpreted as in support of the two-stage models of negation comprehension, while, of course, an opposite result can be interpreted as against the proposal of these models. With respect to affirmative sentences, we are aware of the possibility to find an effect due to the numerosity factor, but we do not have any prediction on the direction of the effect.

Experiment 1

Methods

Participants

We recruited 24 participants. The sample size was exploratory. Participants were recruited through posts on Facebook in one unofficial page of the University of Trento ("*Bacheca Esperimenti Rovereto Trento Mattarello").* All participants were between 18 and 40 years old, Italian native speakers, with no uncorrected deficits in vision and with no neurological or psychological problems. In case they were students of the Psychology Departments they received course credits for their participation.

Stimuli

The experiment was a sentence – picture verification task. We constructed a total of 16 different experimental sentences in Italian, 8 affirmatives and 8 negatives. All the sentences were of the form "I X sono/non sono Y" ("The X are/are not Y"). The noun in the subject position could be "cerchi" ("circles") or "quadrati" ("squares"), while in the adjective position there could be one of four colour terms ("rosso", "blu", "verde", "viola"/ "red", "blue", "green", "violet"). Every picture was composed by a grey (RGB: 120,120,120, 25 cd/m2) background, a black fixation cross in the centre, and two groups of shapes: circles (visual angle 0.21°) and squares (visual angle 0.21°), displaced around the cross. The two groups of shapes had two different numerosity, one group was composed by 6 elements and the other one by 24 elements (ratio 1:4) and had two different colours. In total we used 4 different colours. This means that in each picture there were two of these four colours: green (RGB:16,102,24), red (RGB:160,50,20), blue (RGB:9,86,171) or violet (RGB:123,54,142). All the colours were equiluminant (~14 cd/m2). Each of the 24 possible combinations of shape/colour/numerosity was constructed. In

total we created 1.536 pictures divided in 64 pictures tokens for each of the 24 possible combinations of: 2 shapes, and 2 colours out of 4 (6 combinations) and 2 levels of the numerosity variable. The pictures were constructed using a 6X6 matrix (visual angle 7.5°X11.2°). The shapes appeared displaced randomly in the matrix cells with the only two constraints that not all the shapes of the minority group appeared on the border of the matrix and that no stimulus appeared in the four cells around the fixation cross. Lastly, the centre of each shape was randomly jittered away from the centre of the cell along both horizontal and vertical axes, using a uniform probability distribution of jitter amplitude between 0 and 0.9 of half the axis length.

Design

Combining each sentence (affirmative or negative) with the images in which one of the two shapes (circles or squares) was coloured with the colour adjective mentioned in the sentence, we were able to create two different conditions. A condition in which the image correctly represented the meaning of the sentence (true condition), and a condition in which the image did not represent correctly the sentence meaning (false condition) (cnfr. Table1). In half of the images, the subject of the sentence, circles or squares, were the majority of the shapes depicted, in the other half the minority. This created a 2 (Polarity: affirmative/negative) X 2 (Truth_Value: true/false) X 2 (Numerosity: majority/minority) experimental design. Moreover, since each participant saw each experimental condition, our experiment was a within-subjects design. The experiment was divided in 8 blocks of 32 trials each. Within a block all the levels of the experimental factors were numerically equal represented and randomly presented. In each trial the picture presented was randomly selected between the 64 compatible pictures with that specific trial.

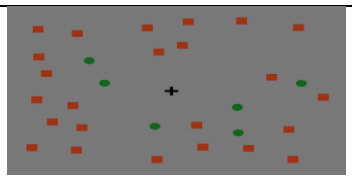
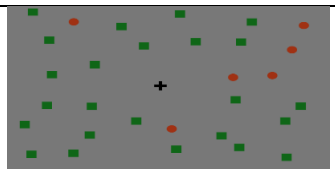
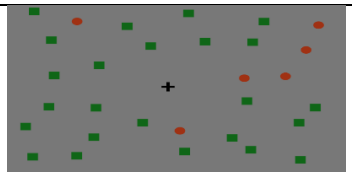
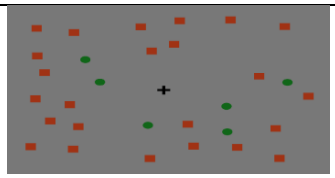
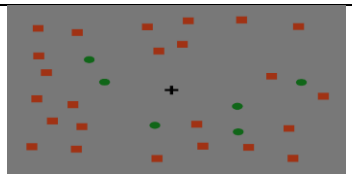
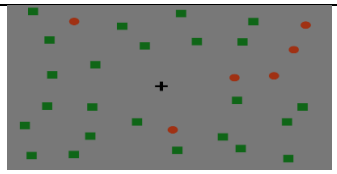
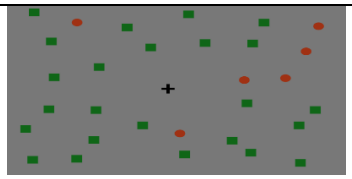
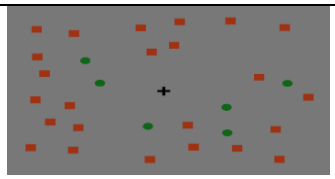
SENTENCE	IMAGE True Condition	IMAGE False Condition
I cerchi sono verdi *The circles are green*		
I cerchi sono rossi *The circles are red*		
I cerchi non sono rossi *The circles are not red*		
I cerchi non sono verdi *The circles are not green*		

Table1. Example of experimental conditions for sentences referring to the minority of objects (circles).

Procedure

Participants were seated in a quiet room with the luminance of the room kept at 5 cd/m2. The computer screen was placed in front of the participants at a distance of 80 cm. To reduce head movements, subjects were asked to use a chin-rest. The materials were displayed using the Software E-Prime 2 (Psychology Software Tools, Pittsburgh, PA). Each trial began with a blank grey monitor for 1000 ms followed by a black fixation cross on a grey background for 500 ms. Then, immediately below the fixation cross, the sentence was presented word by word, with each word presented for 200 ms followed by a screen with only the fixation cross for 150 ms. After the offset of the sentence the fixation cross remained on the screen for 250 ms before the picture was presented (SOA=600ms between the onset of the last word and of the picture). The picture remained on the screen until the subject's response.

Participants were instructed to keep their gaze on the fixation cross both during the sentence and the picture presentation and to indicate as quickly and accurately as possible whether the picture correctly represented or not the sentence meaning, by pressing one of two buttons, labelled as "true" and "false". Half of the participants made "true" responses with their left index finger (pressing the “z” key on a pc keyboard) and "false" responses with their right index finger (pressing the “m” key). For the remaining participants response hands were reversed. Each test session started with a practice block of 10 trials randomly selected between all the possible experimental conditions. During the practice block participants received feedbacks on their responses, no feedback was given on response accuracy during experimental blocks.

Analysis and Results

We analysed reaction times (RTs) and accuracy in the truth-judgment task using the lme4 package (Bates D, Mächler M, Bolker B, Walker S, 2015) in Rstudio software, 1.2.1335 version (RStudio Team, 2018).

Reaction Times

Reaction times of wrong responses and/or slower than 4s were excluded from the analysis. On the reaction times so corrected we fitted eight different Generalized Linear Mixed Models. All the models had an inverse gaussian distribution (Lo and Andrews, 2015) and a random structure composed by random intercepts for participants (see Table 2). We performed a comparison among these models fitted using AIC index as decision criterion.

Models	**AIC**	**R Df**	**R Dv**	**Δ AIC**	**AICw**	**Df**
RTs ~ (1 | Participants)	82154	5548	0.784	1531.	0	3

RTs ~ Polarity + Truth_Value + (1 | Participants)	80906	5546	0.626	283.	0	5
RTs ~ Polarity + Truth_Value + Numerosity + (1 | Participants)	80860	5545	0.620	236.	0	6
RTs ~ Polarity * Truth_Value + (1 | Participants)	80711	5545	0.604	88.1	0	6
RTs ~ Polarity + Truth_Value * Numerosity + (1 | Participants)	80848	5544	0.619	225.	0	7
RTs ~ Polarity * Truth_Value + Numerosity + (1 | Participants)	80649	5544	0.597	26	0	7
RTs ~ Truth_Value + Polarity * Numerosity + (1 | Participants)	80861	5544	0.620	238.	0	7
RTs ~ Polarity * Truth_Value * Numerosity + (1 | Participants)	80623	5541	0.594	0	1	10

Table2. Models' comparison for Reaction Times. AIC = Akaike Index, R Df = Residual Degree of Freedom, R Dv = Residual Deviance, Δ AIC = delta Akaike, AICw = Akaike weight, Df = Degree of Freedom.

The models' comparison among the eight Generalized Linear Mixed Model showed that the multiplicative model was the model which better approximated the data (lower AIC index). In order to test our hypotheses, we conducted paired comparisons on the model selected, p-values were adjusted with Tukey correction for multiple comparisons. Paired comparisons revealed that, for negative sentences in true condition, the model did not predict difference in the reaction times when the sentences referred to the majority or to the minority of the shapes depicted (β = -32.74, SE = 18.2, z.ratio = -1.795, p = 0.4532), while for affirmative sentences in true condition the model predicted faster answers when the sentences referred to the majority vs the minority (β = -98.53, SE = 11.0, z.ratio = -8.967, p < 0.0001). This effect of faster answer with the majority vs the minority was also predicted with negative sentences in the false condition (β = -111.09, SE = 14.5, z.ratio = -7.686, p < 0.0001), but not with affirmative sentences in the false condition (β = 5.18, SE = 13.5, z.ratio = 0.382. p = 0.9999). Moreover, with negative sentences, the model predicted faster answers in false conditions with respect to true ones, both when the sentences were referred to the majority (β = - 167.99, S.E. = 16.9, z.ratio = - 9.952, p <.0001) and, contrary to our hypothesis, to the minority (β = - 89.64, S.E. = 19.8, z.ratio = - 4.538, p <.0001) of the shapes depicted. With affirmative sentences, instead, the model predicted faster answers in true conditions with respect to the false ones, both when

the sentences were referred to the majority (β = -213.74, S.E. = 12.7, z.ratio = -16.884, p <.0001) and to the minority (β =-110.03, S.E. = 12.5, z.ratio = -8.769, p <.0001) of the shapes depicted (see Figure 1).

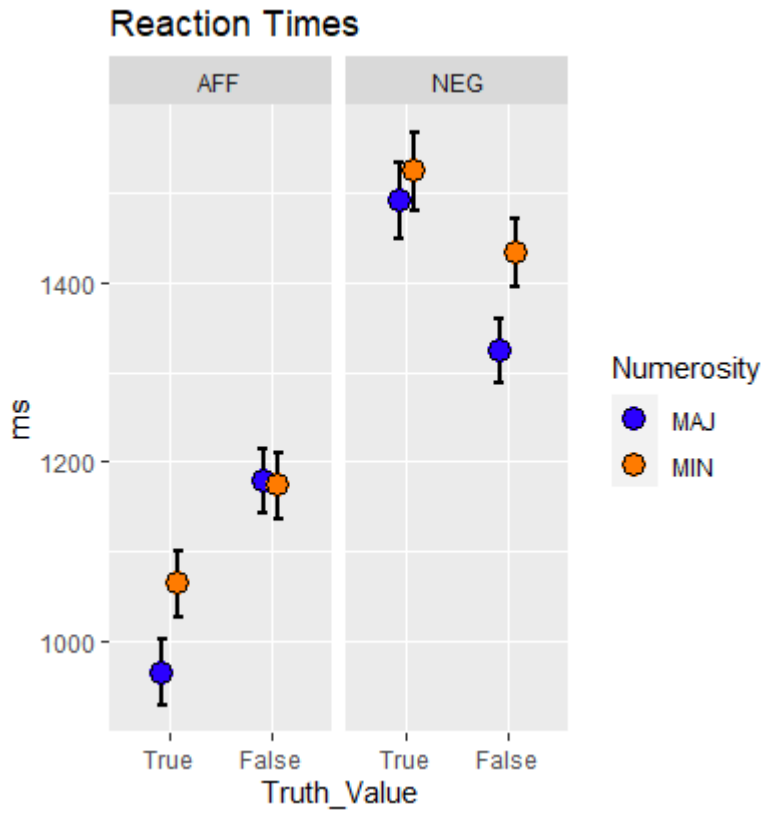


*Figure1. Plot of the reaction times predicted by the model ~ Polarity * Truth_Value * Numerosity + (1 | Participants)*

Accuracy

Responses slower than 4 s were excluded from the analysis. On the remaining responses, we fitted eight Generalized Linear Mixed Effect Models. All the models had a binomial distribution and a random effects structure composed by random intercepts for subjects (see Table 3). We performed a comparison among these models fitted using AIC index as decision criterion.

Models	AIC	R Df	R Dv	Δ AIC	AICw	Df
ACC ~ (1 | Participants)	3594.3	6113	3514.	118	0	2
ACC ~ Polarity + Truth_Value + (1 | Participants)	3488.6	6111	3404.	12.3	0	4
ACC ~ Polarity + Truth_Value + Numerosity + (1 | Participants)	3487.5	6110	3401.	11.2	0	5
ACC ~ Polarity * Truth_Value + (1 | Participants)	3477.4	6110	3391.	1.1	0.35	5
ACC ~ Polarity + Truth_Value * Numerosity + (1 | Participants)	3489.2	6109	3401.	12.9	0	6
ACC ~ Polarity * Truth_Value + Numerosity + (1 | Participants)	3476.3	6109	3388.	0	0.61	6
ACC ~ Truth_Value + Polarity * Numerosity + (1 | Participants)	3489.5	6109	3401.	13.2	0	6
ACC ~ Polarity * Truth_Value * Numerosity + (1 | Participants)	3481.8	6106	3387.	5.5	0.04	9

Table 3. Models' comparison for Accuracy. AIC = Akaike Index, R Df = Residual Degree of Freedom, R Dv = Residual Deviance, Δ AIC = delta Akaike, AICw = Akaike weight, Df = Degree of Freedom

The models' comparison among the Generalized Linear Mixed Effect Models showed that the model with Polarity and Truth_Value interaction and the main effect of Numerosity was the model which better approximated the data (lower AIC index). We conducted paired comparisons on the model, p-values were adjusted with Tukey correction for multiple comparisons. With respect to the main effect of Numerosity, although the AIC index signalled that adding this factor the model better approximated the data, paired comparisons revealed that the model did not predict a significant difference in accuracy depending on the numerosity of the group of shapes which the sentences referred to (Majority – Minority: $\beta = 0.161$, SE = 0.0911, z.ratio = 1.770, p = 0.0767). The interaction effect between Polarity and Truth_Value, instead, showed that the model predicted more accurate responses in affirmative true with respect to affirmative false conditions ($\beta = 0.478$, SE = 0.164, z.ratio = 2.922, p = 0.0070) while the difference in the direction of an higher accuracy in negative false with respect to negative true conditions did not reach significance ($\beta = 0.239$, SE = 0.111, z.ratio = 2.158, p = 0.0609) (see Figure 2).

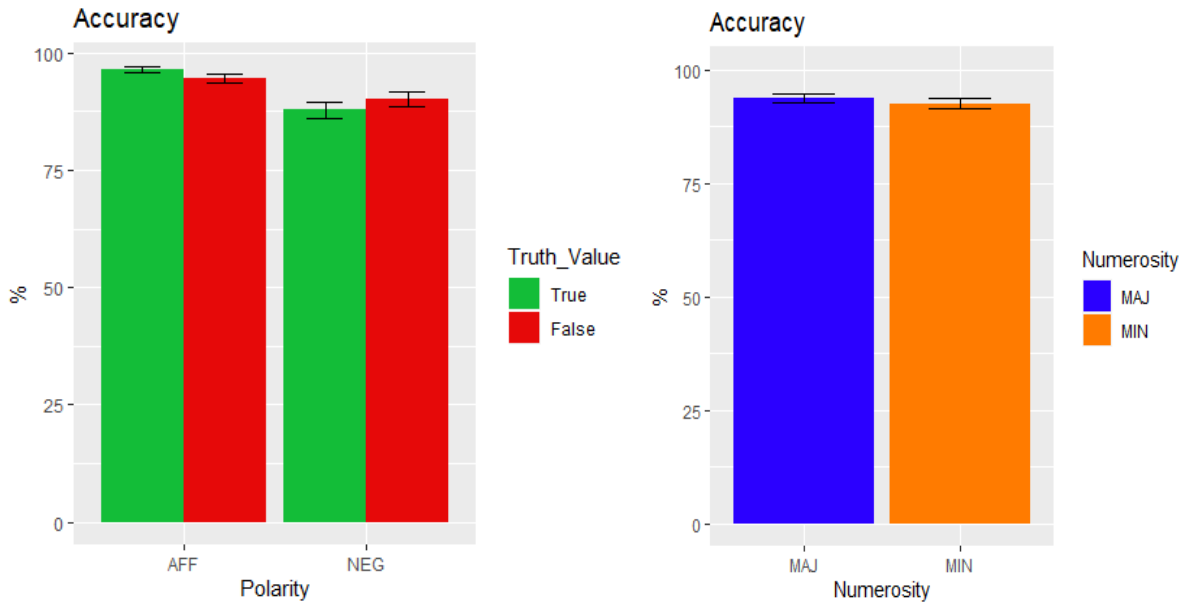


Figure 2. (left) plot of the interaction on Accuracy between Polarity and Truth_Value and (right) plot of the main effect (non-significant) of Numerosity predicted by the model ~ Polarity * Truth_Value + Numerosity + (1 | Participants)

Discussion

The aim of this experiment was to investigate if the fact that a negative sentence refers to a minority is one of those pragmatic factors that allows to access immediately the actual meaning of a negative sentence. To test this proposal, we implemented a sentence-picture verification task in which affirmative and negative sentences described correctly (true condition) or not (false condition) the colour of one group of geometrical shapes depicted in a subsequent picture. In every picture, two different groups of shapes with two different numerosity were represented, with one of the two groups, the majority one, that had twenty-four elements and the minority group that had six elements. The task of participants was to evaluate if the picture correctly represented the sentence meaning or not, responding true or false. We analyzed reaction times and accuracy of participants responses in the verification task. In line with the *ratio hypothesis* proposed by Wason, which foresees a facilitation in the interpretation of negations when referred to a minority, we predicted a reduction of the costs in the negative true

condition when the sentences were referred to the minority of shapes, in comparison to when the sentences were referred to the majority. As stated in the introduction of this experiment, we would have not considered this result sufficient in order to confirm the *ratio hypothesis*, but only if this facilitation with the minority was not found in the affirmative true condition. As a second hypothesis, we predicted that if referring to a minority is a factor that allows to access immediately the actual meaning of a negative sentence, participants would have been facilitated in answering to negative true condition with respect to the negative false one, when the sentences were referred to the minority. Our results, however, do not seem to confirm any of our hypotheses. Indeed, with respect to our first hypothesis, our results indicated no difference in the negative true condition between sentences referred to the majority or to the minority. Moreover, we found slower reaction times in the affirmative true condition when sentences were referred to the minority of the shapes. We think that this result cannot be interpreted as evidence of an easier interpretation of negative sentences referred to the minority, and so it seems to be at odds with Wason's *ratio hypothesis*. At the same time, following the reasoning of Wason (1965) presented in the Introduction of this chapter, one can speculate that we failed to confirm the *ratio hypothesis* for the same reason as Wason's original failure. Despite we think that the random and intermixed presentation of shapes of different colour in the screen should be conceptualized as a "single set" we cannot categorically exclude participants coded the images presented as composed by two distinct groups of geometrical shapes, one group of circles of one colour, and one group of squares of another colour. In light of this, we cannot exclude that in a paradigm which would force (even more than we tried) participants to code the images as composed by one unique group of geometrical shapes, the *ratio hypothesis* could be confirmed. One possible idea to try to further force participants to code the images as composed by one unique group of items would be that to present images in which only one type of shapes (e.g. circles) are presented, with a majority of the shapes coloured with one

colour and a minority of them coloured with another colour. Alternatively, one could present pictures in which all items have the same colour, with a majority of items sharing the same shape (e.g. circles) and a minority of them having another shape (e.g. squares). Of course, the use of this type of images would rise the problem of the construction of sentences that are true or false. One idea to overcome this issue could be that to use sentences with quantifiers of the type "The majority of the X are/are not Y" vs "The minority of X are/are not Y". However, this manipulation should be carefully implemented since a great debate concern if positive (e.g. "the majority") and negative (e.g. "the minority") quantifiers are comprehended in different ways (see for example Urbach, Kutas; 2010). Anyway, moving to our second hypothesis, and so the hypothesis of the immediate access to the actual meaning of the negation for the minority, it is not confirmed because, when the negative sentences were referred to both the majority and the minority, our results showed faster reaction times in the false condition with respect to the true one. Of course, this result can be interpreted as supporting the proposal of the two-stage models of negation comprehension, and so as evidence that even when the negative sentences are referred to the minority, participants access the affirmative counterpart meaning of the sentence. The effect of facilitation in the false vs true condition with negative sentences is usually considered sufficient to assume an early access to the affirmative counterpart of the sentence meaning and thus a two-stages processing. Anyway, we think that, in our experiment, additional evidence in support of this interpretation comes from the numerosity effect we found on reaction times. Indeed, our results showed faster answers for affirmative sentences when these were referred to the majority with respect to the minority, but this happened only in the true condition and not in the false condition. This result indicates that when after an affirmative sentence, as for example "The circles are red", a picture was presented in which red circles and green squares were depicted, participants were faster in responding when the red circles were the majority rather than the minority of the items

depicted. However, the numerosity effect disappeared when green circles and red squares were depicted in the picture presented. Of course, from our data, it is impossible, to explain the reason of this effect. One can speculate that the effect is explainable by the fact that participants, before responding, checked both features named in the sentence. Thus, in the true condition, being the two features (red and circles) presented jointed on the majority or on the minority of items, participants were faster to check the two features in the majority condition because it is easier to identify the items if they are the majority rather than the minority. In the false condition, instead, being the two features (red and circles) presented disjointed, one on the majority and the other one on the minority of items, the numerosity effect disappeared because, in this case, participants had to check both the items. However, this is just one of the possible explanations, future research is needed to explain the effect.

Be as it may, it is interesting to note that with negative sentences we found exactly the specular result we found with affirmative sentences: faster reaction times in the majority condition in the false condition and not in the true condition. This means that also after a negative sentence, as for example "The circles are not red", when the sentence was followed by a picture with red circles and green squares, participants were faster when the red circles were the majority rather than the minority of the items depicted. However, the numerosity effect disappeared when green circles and red squares were depicted in the picture. We think that this specular result can be seen as a nice confirmation of the idea that in performing our task participants treated negative sentences exactly as if they were their affirmative counterpart. So, we think that this effect can be interpreted as a nice result in support of the idea of the two stage models of negation comprehension and so as evidence of the fact that in the interpretation of the negation participants pass through a moment in which access the affirmative counterpart meaning of the sentence.

Chapter 4: Nonexistence[3]

Introduction

A negative sentence can be used to express at least two things: nonexistence or an alternative. For example, a sentence like "In this box there are no candies" might mean both that in the box there is nothing and that in the box there are other things, such as for examples chocolates (Nordmeyer & Frank; 2015). In the first case negation is used to express the nonexistence of candies in the box while, in the second one, is used to describe the presence of an alternative to candies. In the literature, two experiments by Nordmeyer and Frank (2014, 2015) seem to suggest an easy processing and pragmatic felicity for negations that expresses nonexistence with respect to negations that expresses an alternative. Nordmeyer and Frank (2014) discuss the observation that at developmental level the production of negation to express nonexistence appears earlier than that to express an alternative (Pea, 1980). On the basis of this observation, they decide to investigate how children process non-existence negation and alternative negation in two separate preferential looking experiments using eye-tracking. For both experiments the authors recruited three groups of children between 2 and 5 years old, and one group of adults as control group. In the first experiment participants saw a picture with two characters, one character holding two objects (e.g. apples), and one character holding nothing. During the presentation of this picture a recorded voice said an affirmative or a negative sentence, as: "Look at the character with/with no apples. Can you find him?". The authors registered the looks of participants toward the target character from the onset of the critical word (e.g. apples). In the second experiment the authors used exactly the same paradigm with the only difference that the mismatching character (the one without apples) held two alternative objects (e.g.

[3]These studies have been carried out in collaboration with Chiara Finocchiaro and Francesco Vespignani

boxes) instead of nothing. The results of the children's groups are pretty complex and not homogeneous across groups. However, being these results not crucial for our work, we will limit to report that, according to the authors' interpretation, children found more difficulties in the processing of the negation in the first experiment (nonexistence negation) with respect to the one in the second experiment (alternative negation) (for a detailed explanation of the results we would send back you to the paper by Nordmeyer and Frank (2014)). More interesting for our work, are the results of the group of adults. Indeed, the group of adults seemed to be facilitated in the processing of the nonexistence negation vs the alternative negation, as evidenced by the fact that they were faster to move their gaze toward the right target with the nonexistence negation. Moreover, in the first experiment (nonexistence negation), adults were equally fast to identify the target with the negative and the affirmative sentence while in the second one (alternative negation) they were faster to identify it with the affirmative sentence. A result of this type could be interpreted as evidence of a processing of the nonexistence negation that is as easy as the processing of an affirmative sentence, while the processing of the alternative negation is more difficult than the processing of an affirmative sentence. However, there is an issue in the nonexistence experiment that makes difficult to disentangle if the effect found with adults was due to a facilitated processing of the negation or simply to anticipation. Indeed, given that in the picture presented in this experiment were represented one character that was holding two objects and one character that was holding nothing, when presented with a negative sentence such as "Look at the character with no apples. (…) ", participants could predict that the right target would have been the one that was holding nothing as soon as they heard the word "no", and so before to hear the critical word (e.g. "apples"). Differently, with the affirmative sentence (e.g. "Look at the boy with apples. (…) ") and in the experiment that tested the alternative negation, in which was presented a picture in which both the characters held two objects (e.g. apples and boxes), participants had necessarily to hear the

critical word (e.g. “apples”) before the identification of the right target. In a following study Nordmeyer & Frank (2015) studied whether nonexistence negation and alternative negation differ in terms of pragmatic felicity by presenting a group of adults’ participants with pictures in which four cartoon characters were depicted. In some pictures the target character was holding nothing, while in others he/she was holding an object (e.g. apple). All the pictures were paired with a sentence that might be affirmative or negative and might describe correctly or not what the target character was, or was not, holding (e.g. “Abby have/doesn’t have a banana”). The task of participants was to rate the sentence-picture match on a seven-point scale going from “Very Bad” to “Very Good”. When the sentence correctly described the picture, participants rated the negative sentences as more apt to describe the situation when the character was holding nothing with respect to the situation in which he/she was holding an object. Based on this result the authors concluded that a negative sentence is more felicitous if it is used for describing nonexistence rather than an alternative.

Given the results of Nordmeyer and Frank works (2014, 2015) that possibly suggest an easier processing but, more importantly, a stronger pragmatic felicity for the nonexistence negation we think it is worth to investigate if using a negation to describe something that non-exists is one of those pragmatic factors that allows to access immediately the actual meaning of a negative sentence. To do this, we implemented two sentence-picture verification experiments.

Experiment 2

In this experiment we compared negative sentences used to describe something that non-exists versus an alternative. To do this we presented two negative sentences that differed in the type of adjective. Indeed, for the nonexistence negation, we presented sentences as “The circles/squares are not coloured” while for the alternative negation we presented sentences as “The circles/squares are not black”. For the nonexistence negation, the sentence “The

circles/squares are not coloured", was followed by a picture composed by two groups of shapes (circles and squares) in which one of the two groups was coloured, using different colours, and the other one was not coloured, being filled with the same grey tone as the picture background (see Table 1 for pictures examples). For the alternative condition, the sentence, "The circles/squares are not black", was followed by a picture in which one of the two groups was coloured in black and the other one in white. Since in the nonexistence condition participants could predict that "not coloured" meant grey as the picture background, we created an alternative condition in which participants could predict that "not black" meant white. We decided to use the white colour because in natural language use white is the opposite of black. The task of participants was to evaluate if the picture presented after the sentence represented correctly the sentence meaning by choosing between "true" and "false" responses. With respect to the nonexistence negation, we predicted that if the fact to describe the nonexistence allows to access immediately the actual meaning of the negative sentence, participants would have been faster and/or more accurate in responding in true conditions with respect to false ones. For the alternative negation, instead, we predicted to find the classical result, that is often interpreted as evidence of an access to the affirmative counterpart meaning of the sentence, for which participants are faster and/or more accurate in false conditions with respect to true ones. The two conditions were presented to all subjects in a within participant design, however we decided to keep separated the presentation of the nonexistence negation and of the alternative negation. So, in four experimental blocks we presented only sentences with the coloured adjective, and in other four blocks, only sentences with the black adjective. Within the experiment we presented initially all the four blocks for one adjective and after all the four blocks for the other adjective. In the respective blocks, we also tested the affirmative counterpart of the negative sentences and so "The circles/squares are coloured" and "The

circles/squares are black". For both affirmative sentences, we predicted to find faster and or/more accurate answers in true with respect to false conditions.

Methods

Participants

We planned to recruit fifty participants for this experiment. The sample size was exploratory. Due to the covid pandemic we decided to run this experiment online and to implement it in English for easier recruitment. Participants were recruited via Prolific system (www.prolific.co), for their participation they have been rewarded with 3.5£. All participants declared to be between eighteen and forty years old, to be English mother-tongue, to have at least a secondary level education, to not have literacy difficulties, to have normal or corrected to normal vision, to not take any drugs for mental illness currently and to not have or have had diagnosis of mental illness. In total 53 participants completed the experiment but three of them have been removed from the sample because of technical problems (two of them) and low effort in the task (one participant preformed with an overall accuracy lower than 0.8)).

Stimuli

The experiment was a sentence-picture verification task. The sentences were those described two sections above (see Table 1 for an example).

One hundred-and-twenty-eight pictures were constructed offline as png images in order to simplify the programming of the presentation software (Jatos, see below) and to control details of the visual display. All the images had dimension 800x600 pixels and were composed by a grey background (rgb=120, 120, 120) and by 6 circles (diameter = 24 pixels) and 6 squares (side = 24 pixels), mixed between them, displaced around a black fixation cross (side=24 pixels) positioned in the centre of the image. The pictures were constructed using a

4(rows)X6(columns) matrix; every cell of the matrix had dimension of 60x60 pixels. Circles and squares could appear in every matrix cell with random jitter, the only constraint was that the four cells around the fixation cross were empty. The 128 total pictures were divided in four different type: 32 pictures in which circles were coloured and squares were not, 32 pictures in which squares were coloured and circles were not, 32 pictures in which circles were black and squares were white, 32 pictures in which squares were black and circles were white. For the 64 images in which the shapes could be coloured or not, circles and squares had the edge. For the coloured shapes a maximum of four items could be presented, on the total of six, differently coloured simultaneously (red (rgb=160,50,20); green (rgb=16,102,24); blue (rgb=9,86,171); violet (rgb=123,54,142)), for the not coloured shapes only the edge was drawn. For the 64 images in which the shapes could be black or white, instead, circles and squares did not have the edge. Random examples of the pictures in the four conditions are reported in Table 1, All the stimuli can be retried from https://osf.io/c5yds/.

Design

As stated above, our experimental sentences (affirmative or negative) could have one of two adjectives, coloured or black. The sentences with the coloured adjective have been paired with the images in which one of the two shapes depicted, circles or squares, were coloured, while the sentences with the black adjective with the images in which one of the two shapes depicted, circles or squares, were black. The images could correctly represent the meaning of the sentence (true condition) or not (false condition) (cnfr. Table 1). This created a 2 (Polarity: affirmative/negative) X 2 (Truth_Value: true/false) X 2 (adjective of the sentence: coloured/black) experimental design. Moreover, since each participant saw each experimental condition, our experiment was a within-subjects design.

The experiment was divided in 8 blocks of 32 trials each. Within a block only one of the two levels of the adjective factor was presented (coloured or black) while all the remaining levels of the experimental factors were numerically equal represented and randomly presented.

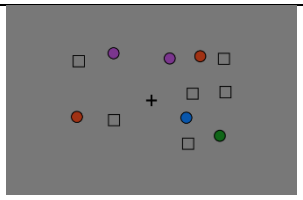
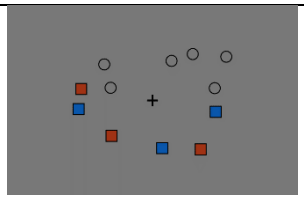
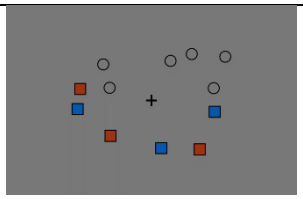
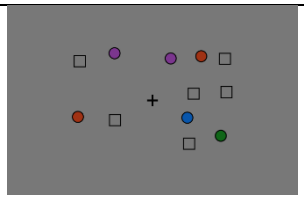
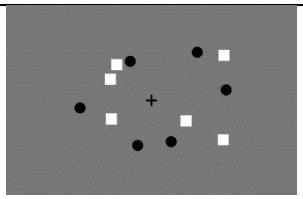
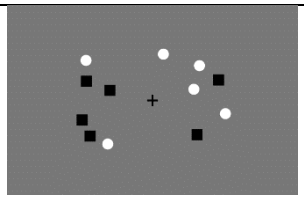
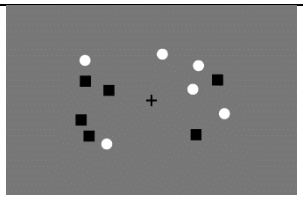
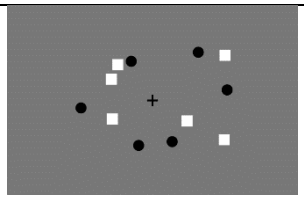
SENTENCE	IMAGE True Condition	IMAGE False Condition
The circles are coloured		
The circles are not coloured		
The circles are black		
The circles are not black		

Table 1. Example of the experimental conditions.

Procedure

Participants were informed that the experiment worked only on desktop-computer (not on tablets or smartphones). After selecting the experiment in Prolific, participants were redirected to a server running the Jatos software (Lange, Kühn, & Filevich,. 2015). The material of the experiment was displayed using the OpenSesame On-line Software (Mathôt, Schreij, & Theeuwes, 2012). Participants read the instructions about how to set up the environment to

participate in the experiment. These instructions asked them to find a quiet room where they could complete the experiment without any distractions, to place their computer on a table in front of them and to sit on a backrest chair. Moreover, they asked them to close all other applications and browser windows other than the one running experiment and set its window full screen. Lastly, they asked them to check that on their computer screens there were not halos or light reflections that could make reading difficult. After these general instructions about setup the experimental instructions were presented. Participants were instructed that in each trial a sentence will be presented word by word and will be followed by an image. After the image disappearance they were asked, to respond as faster and more accurately as possible "true" if the image correctly represented the meaning of the sentence and "false" in the other case, by using the "z" and "m" keys of their keyboard. A random selection determined if participant had to respond true with the "z" key and false with the "m" key or the opposite. Moreover, participants were randomly assigned to one of two different groups: for one group the first four experimental blocks presented the coloured adjective condition and the last four presented the black adjective condition, for the other group the first four experimental blocks presented the black adjective condition and the last four presented the coloured adjective condition. Independently from the group assigned, all participants, before the experiment, completed a practice block composed by 12 trials in which all the levels of the experimental factors (and so both the adjective factor levels, coloured and black) were presented intermixed in a random order. Both in the practice block and in the 8 experimental blocks each trial began with a dark fixation cross in the centre of the grey screen (rgb = 120, 120, 120) for 1 second. Then the sentence was presented word by word 32 pixels under the fixation cross, with each word presented for 300ms and with an ISI between words of 200ms. During the ISI only the fixation cross in the centre of the screen was presented. With a SOA of 600ms from the onset of the last word of the sentence (adjective word = 300ms + ISI = 300ms) the image was

presented for 250ms, and subsequently a blank screen was presented until the participant responded. In the practice block, after each response, participants received feedback on the correctness of their response, a green "CORRECT" feedback in the case of a correct answer and a red "WRONG" feedback in the case of a wrong answer, instead this did not happen in the experimental blocks in which, immediately after the participant response, and without any feedback, the following trial started.

Analysis and Results

We analysed reaction times (RTs) and accuracy in the truth-judgment task using the lme4 package (Bates, Mächler, Bolker, Walker, 2015) in Rstudio software, 1.2.1335 version (RStudio Team, 2018).

Reaction Times

Reaction times for wrong responses and/or slower than 4s were excluded from the analysis. Reaction times so corrected were normalized using the logarithmic transformation in base 10 and on these we fitted 8 different Linear Mixed Effect Models with a random effects structure composed by random intercepts for participants (see Table 2).

Models	**AIC**	**R Df**	**R Dv**	**Δ AIC**	**AICw**	**Df**
RTs ~ (1 | Participants)	230.87	11850	225.	1500.	0	3
RTs ~ Polarity + Truth_Value + (1 | Participants)	-1220.93	11848	-1231.	48	0	5
RTs ~ Polarity + Truth_Value + adjective + (1 | Participants)	-1225.43	11847	-1237.	43.5	0	6
RTs ~ Polarity * Truth_Value + (1 | Participants)	-1238.38	11847	-1250.	30.5	0	6
RTs ~ Polarity + Truth_Value * adjective + (1 | Participants)	-1223.78	11846	-1238.	45.1	0	7
RTs ~ Polarity * Truth_Value + adjective + (1 | Participants)	-1242.69	11846	-1257.	26.2	0	7
RTs ~ Truth_Value + Polarity * adjective + (1 | Participants)	-1247.72	11846	-1262.	21.2	0	7
RTs ~ Polarity * Truth_Value * adjective + (1 | Participants)	-1268.91	11843	-1289.	0	1	10

Table 2. Models' comparison for Reaction Times. AIC = Akaike Index, R Df = Residual Degree of Freedom, R Dv = Residual Deviance, Δ AIC = delta Akaike, AICw = Akaike weight, Df = Degree of Freedom.

The models' comparison among Linear Mixed Effect Models showed that the multiplicative model among Polarity, Truth_Value and adjective was the model which better approximated the data (lower AIC index). In order to test our hypotheses, we conducted planned comparisons on the model selected, p-values were adjusted with Tukey correction for multiple comparisons. Within planned comparisons we decided to not compare directly the two levels of the adjective factor (black vs. coloured) because of the difference in the visual complexity of the images presented in the two conditions. Indeed, the difference in the visual complexity of the pictures can justify differences on the dependent variable, making it impossible to disentangle if the observed differences were due simply to the difference in the visual complexity of the images or to the experimental manipulation. Planned comparisons revealed that the model predicted faster answers in true with respect to false condition with affirmative sentences, both with black (β = -0.1317, S.E. = 0.00824, z.ratio = -15.997, $p < .0001$) and coloured (β = -0.1049, S.E. = 0.00829, z.ratio = -12.657, $p < .0001$) adjective. The same result was predicted with negative sentences as well, both with black (β = -0.0722, S.E. = 0.00844, z.ratio = -8.555, $p < .0001$) and coloured (β = -0.0910, S.E. = 0.00839, z.ratio = -10.848, $p < .0001$) adjective (see Figure 1).

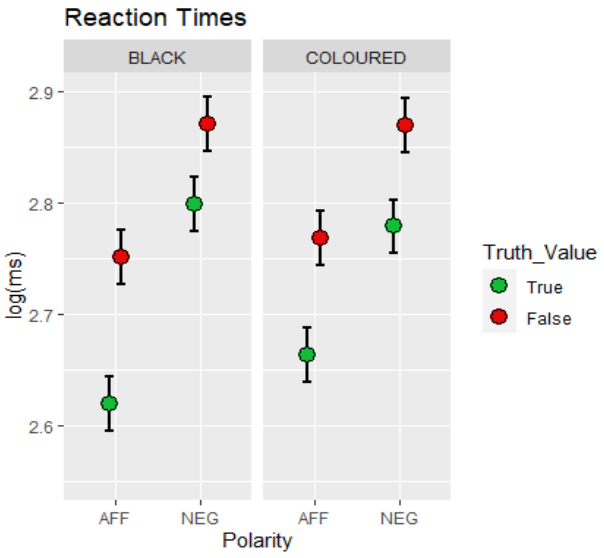


*Figure 1. Plot of the predicted log-trasformed Reaction Times by the model Polarity*Truth_Value*adjective + (1|Participants).*

Accuracy

Responses slower than 4 s were excluded from the analysis. On the remaining responses we fitted 8 different Generalized Linear Mixed Effect Models. The distribution used was a binomial distribution and the random effects' structure was composed by random intercepts for participants (see Table 3)

Models	AIC	R Df	R Dv	Δ AIC	AICw	Df
Acc ~ (1 | Participants)	6092.8	12715	5945.	55.8	0	2
Acc ~ Polarity + Truth_Value + (1 | Participants)	6056.1	12713	5904.	19.1	0	4
Acc ~ Polarity + Truth_Value + adjective + (1 | Participants)	6057.9	12712	5904.	20.9	0	5
Acc ~ Polarity * Truth_Value + (1 | Participants)	6044.8	12712	5891.	7.8	0.02	5
Acc ~ Polarity + Truth_Value * adjective + (1 | Participants)	6059.3	12711	5903.	22.3	0	6
Acc ~ Polarity * Truth_Value + adjective + (1 | Participants)	6046.6	12711	5891.	9.6	0.01	6
Acc ~ Truth_Value + Polarity * adjective + (1 | Participants)	6054.7	12711	5899.	17.7	0	6
Acc ~ Polarity * Truth_Value * adjective + (1 | Participants)	6037.0	12708	5875.	0	0.97	9

Table 3. Models' comparison for Accuracy. AIC = Akaike Index, R Df = Residual Degree of Freedom, R Dv = Residual Deviance, Δ AIC = delta Akaike, AICw = Akaike weight, Df = Degree of Freedom

The models' comparison among Generalized Linear Mixed Effect Models showed that the multiplicative model among Polarity, Truth_Value and adjective was the model which better approximated the data (lower AIC index). In order to test our hypotheses, we conducted paired comparisons on the model selected, p-values were adjusted with Tukey correction for multiple comparisons. For the same reason explained in the reaction times analysis section, also in this analysis, within paired comparisons, we decided to not directly compare the two adjective factor levels (black vs coloured). Paired comparisons revealed that the model predicted more accurate answers in true condition with respect to false one with affirmative sentences for the black adjective ($\beta = 0.674$, S.E. = 0.172, z.ratio = 3.913, $p = 0.0004$), but no difference for the coloured adjective ($\beta = 0.190$, S.E. = 0.150, z.ratio = 1.262, $p = 0.6042$). With negative sentences, the model predicted more accurate answers in false condition with respect to true

one for the black adjective (β = 0.350, S.E. = 0.129, z.ratio = 2.707, p = 0.0268), but no difference for the coloured adjective (β = - 0.105, S.E. = 0.134, z.ratio = - 0.787, p = 0.8955) (see Figure 2).

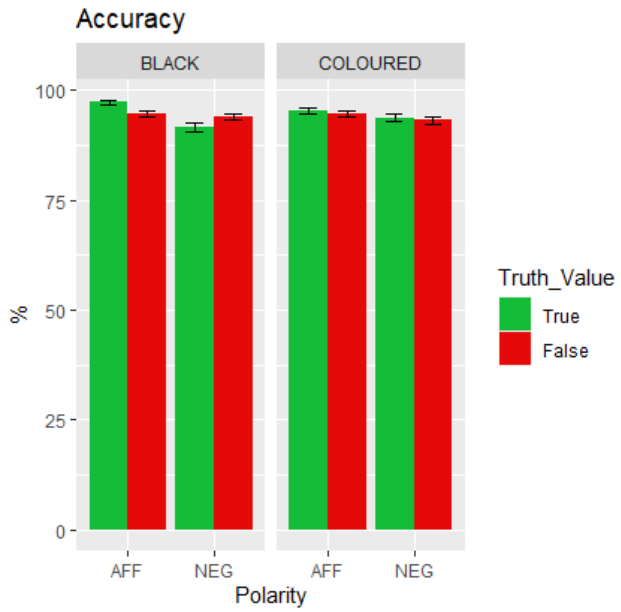


*Figure 2. Plot of the predicted accuracies by the model Polarity*Truth_Value*adjective + (1|Participant).*

Discussion

In this experiment, within a sentence-picture verification task, we investigated if describing something that non-exists is one of those pragmatic factors that allows to access immediately the actual meaning of a negative sentence. To do this, we compared one negative sentence used to describe nonexistence ("The circles/squares are not coloured") with one used to describe an alternative ("The circles/squares are not black"). We predicted that, if describing nonexistence allows to access immediately the actual meaning of a negative sentence, participants would have been faster and/or more accurate in responding in true condition with respect to false one when negation was used to describe nonexistence. Instead, when negation was used to describe an alternative, we predicted to find the classical result for which participants are faster and/or more accurate in responding in false condition with respect to true one. For the nonexistence negation the results are pretty clear. Indeed, in line with our hypothesis, our results showed

faster answers in the true condition with respect to the false one, while no difference between true and false condition was predicted on the accuracy. This result is evidence of an easier task in the true rather than false condition, and so could be interpreted as a sign of an occurred access to the actual meaning of the negative sentence in the case in which negation was used to express nonexistence. For the alternative negation, instead, the results are less clear. Indeed, while the results on accuracy are in line with our hypothesis, and so higher accuracy in false condition with respect to true one, results on reaction times showed exactly the opposite pattern and so faster answers in the true condition with respect to the false one. Of course, this speed-accuracy trade off effect does not allow us to make any conclusion about the confirmation or rejection of our hypothesis.

Experiment 3

In Experiment 2 we found evidence for an occurred access to the actual meaning of the negative sentence that described nonexistence ("The circles/squares are not coloured") and a speed-accuracy trade off that does not allow any definitive conclusion with respect to the negation that described an alternative ("The circles/squares are not black"). Anyway, one methodological aspect of Experiment 2 could have affected the results. Indeed, in the first experiment, within each experimental block we presented only one of the two negations ("The circles/squares are not coloured" or "The circles/squares are not black"). This means that within one block only one adjective (coloured or black) was presented. As a consequence, within each experimental block, participants could fully predict what would have been the negated word. In this experiment 3 we investigated if the results we found in experiment 2 were affected by the possibility to fully predict the negated word. To do this, we implemented an experiment as the previous one with the only difference that within each experimental block both types of negations, and so both adjectives (black and coloured), were presented. Of course, in this way

participants could not fully predict what would have been the word negated within each block, but only with 50% of accuracy.

Methods

Participants

As in experiment 2, we planned to recruit fifty participants. They were recruited via Prolific system (www.prolific.co) and they were rewarded with 3.5£ for their participation. All participants declared to be between eighteen and forty years old, to be English mother-tongue, to have at least a secondary level education, to not have literacy difficulties, to not have any language impairment, to have normal or corrected to normal vision, to not take any drugs for mental illness currently and to not have or have had diagnosis of mental illness. In total 58 participants completed the experiment but eight of them were removed from the sample because of low effort in the task (they performed with an overall accuracy lower than 0.8). All the 58 participants did not take part in Experiment 2.

Stimuli

The Stimuli were those used in Experiment 2.

Design

The Design of the experiment was the same as in Experiment 2 with the only difference that both levels of the adjective factor (coloured and black) were presented within an experimental block. So, within each block, all the levels of the experimental factors were numerically equally represented and randomly presented.

Procedure

The procedure followed in this experiment was the same followed in Experiment 2.

Analysis and Results

Reaction Times

Reaction times of wrong responses and/or slower than 4s were excluded from the analysis. Reaction times so corrected were normalized using the logarithmic transformation in base 10 and on these we fitted 8 different Linear Mixed Effect Models with a random effects structure composed by random intercepts for participants (see Table 4).

Models	AIC	R Df	R Dv	Δ AIC	AICw	Df
RTs ~ (1 | Participants)	35.29	11912	29.3	2470	0	3
RTs ~ Polarity + Truth_Value + (1 | Participants)	-2271.62	11910	-2282.	163.	0	5
RTs ~ Polarity + Truth_Value + adjective + (1 | Participants)	-2274.13	11909	-2286.	161.	0	6
RTs ~ Polarity * Truth_Value + (1 | Participants)	-2346.51	11909	-2359.	88.2	0	6
RTs ~ Polarity + Truth_Value * adjective + (1 | Participants)	-2272.46	11908	-2286.	162.	0	7
RTs ~ Polarity * Truth_Value + adjective + (1 | Participants)	-2348.91	11908	-2363.	85.8	0	7
RTs ~ Truth_Value + Polarity * adjective + (1 | Participants)	-2333.87	11908	-2348.	101.	0	7
RTs ~ Polarity * Truth_Value * adjective + (1 | Participants)	-2434.68	11905	-2455.	0	1	10

Table 4. *Models' comparison for Reaction Times. AIC = Akaike Index, R Df = Residual Degree of Freedom, R Dv = Residual Deviance, Δ AIC = delta Akaike, AICw = Akaike weight, Df = Degree of Freedom.*

The models' comparison among Linear Mixed Effect Models showed that the multiplicative model among Polarity, Truth_Value and adjective was the model which better approximated the data (lower AIC index). We conducted planned comparisons on the model selected, p-values were adjusted with Tukey correction for multiple comparisons. Planned comparisons revealed that the model predicted faster answers in true with respect to false condition with affirmative sentences, both with the black (β = -0.1394, S.E. = 0.00780, z.ratio = -17.863, $p < .0001$) and the coloured (β =-0.1038, S.E. = 0.00786, z.ratio = -13.212, $p < .0001$) adjective. The same result was predicted with negative sentences as well, both with black (β = -0.0268, S.E. = 0.00805, z.ratio =-3.332, p = 0.0034) and coloured (β = -0.0761, S.E. = 0.00800, z.ratio = -9.514, $p < .0001$) adjective (see Figure 3).

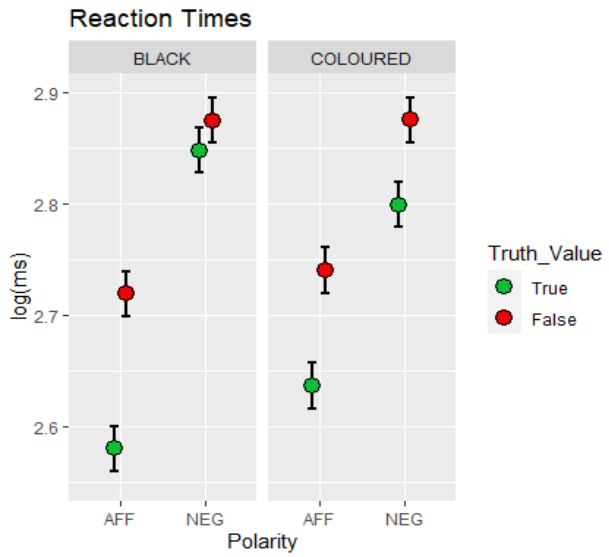


*Figure 3. Plot of the predicted log-trasformed Reaction Times by the model Polarity*Truth_Value*adjective + (1|Participants).*

Accuracy

Responses slower than 4 s were excluded from the analysis. On the remaining responses we fitted 8 different Generalized Linear Mixed Effect Models. The distribution used was a binomial distribution and the random effects' structure was composed by random intercepts for participants (see Table 5)

Models	AIC	R Df	R Dv	Δ AIC	AICw	Df
Acc ~ (1 | Participants)	5832.5	12717	5701.	133.	0	2
Acc ~ Polarity + Truth_Value + (1 | Participants)	5737.7	12715	5602.	38.5	0	4
Acc ~ Polarity + Truth_Value + adjective + (1 | Participants)	5739.7	12714	5602.	40.4	0	5
Acc ~ Polarity * Truth_Value + (1 | Participants)	5700.8	12714	5563.	1.6	0.28	5
Acc ~ Polarity + Truth_Value * adjective + (1 | Participants)	5740.0	12713	5600.	40.7	0	6
Acc ~ Polarity * Truth_Value + adjective + (1 | Participants)	5702.8	12713	5562.	3.5	0.1	6
Acc ~ Truth_Value + Polarity * adjective + (1 | Participants)	5733.2	12713	5593.	34	0	6
Acc ~ Polarity * Truth_Value * adjective + (1 | Participants)	5699.3	12710	5553.	0	0.61	9

Table 5. Models' comparison for Accuracy. AIC = Akaike Index, R Df = Residual Degree of Freedom, R Dv = Residual Deviance, Δ AIC = delta Akaike, AICw = Akaike weight, Df = Degree of Freedom

The models' comparison among Generalized Linear Mixed Effect Models showed that the multiplicative model among Polarity, Truth_Value and adjective was the model which better approximated the data (lower AIC index). We conducted paired comparisons on the model

selected, p-values were adjusted with Tukey correction for multiple comparisons. Paired comparisons revealed that the model predicted more accurate answers in true condition with respect to false one with affirmative sentences for the black adjective (β = 0.7325, S.E. = 0.199, z.ratio = 3.684, p = 0.0009) and for the coloured adjective (β = 0.7015, S.E. = 0.174, z.ratio = 4.024, p = 0.0002). With negative sentences, the model predicted more accurate answers in false condition with respect to true one for the black adjective (β = 0.3615, S.E. = 0.126, z.ratio = 2.867, p = 0.0165), but no difference for the coloured adjective (β = 0.1742, S.E. = 0.134, z.ratio = 1.303, p = 0.5749) (see Figure 4).

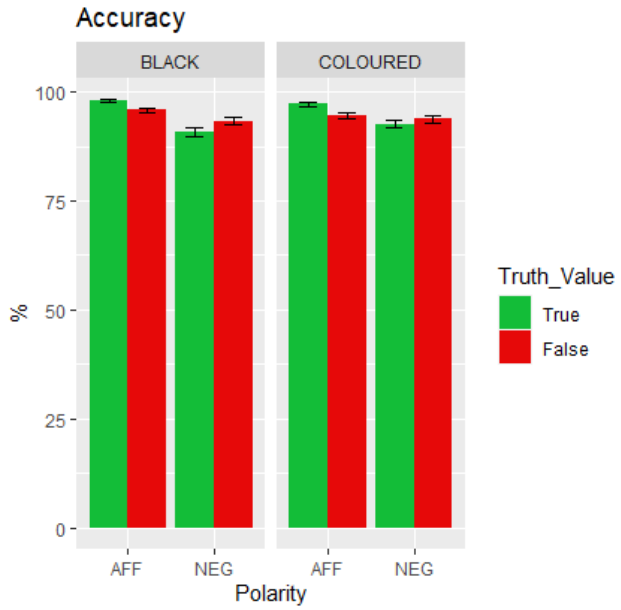


*Figure 4. Plot of the predicted accuracies by the model Polarity*Truth_Value*adjective + (1|Participants).*

Discussion

The aim of experiment 3 was to investigate if the results we found in experiment 2 were affected by the possibility to fully predict the negated word. To do this, in experiment 3, within each experimental block we intermixed the presentation of the two types of negative sentences that in experiment 2 were presented in different blocks ("The circles/squares are not coloured" and "The circle/squares are not black"). This manipulation was operationally translated in the

presentation, within each experimental block, of both the levels of the adjective factor (coloured vs black). In this way, within each experiment block, participants could predict the negated word only with 50% of the accuracy. The results of experiment 3 replicated the results of experiment 2, suggesting that the results of experiment 2 were not strongly affected by the possibility to fully predict the negated word. The only significant difference between the two experiments regards the reached significance in experiment 3 of a higher accuracy in true with respect to false condition with affirmative sentences, when the adjective of the sentences was "coloured". Looking at the accuracy means in the two experiments for the affirmative sentences with the coloured adjective (Exp2: Aff_True = 95,44% vs Aff_False = 94,45%; Exp3: Aff_True = 97,23% vs Aff_False: 94,57%) one can see how this difference is explained by a higher accuracy in affirmative true condition in the second with respect to the first experiment. One could speculate that maybe the greater variety of experiment 3, in which both adjectives were presented within each block, helped participants to stay a little bit more focused on the task, facilitating their performance on the affirmative true sentences when the adjective was coloured.

General Discussion

In line with the idea that negation is more felicitous in expressing nonexistence than alternative (Nordmeyer and Frank, 2015), within two sentence-picture verification experiments we investigated if the fact a negative sentence is used to talk about something that non-exists is one of those pragmatic factors that allows to access immediately its actual meaning. The only difference between the two experiments was that, in the first one, participants could fully predict what would have been the word negated presented in the sentence while, in the second one, only with 50% of accuracy. In both experiments, however, participants could fully predict the affirmative meaning of the negative sentences. We analyzed reaction times and accuracy

of participants' responses in the truth judgment task. For the nonexistence negation we found faster answers in true conditions with respect to false ones, and no difference on the accuracies. This result was consistent across both experiments, indicating that it was not limited to the situation in which it was possible to fully predict the word that would have been negated. This result seems to indicate that with the nonexistence negation we found evidence of the immediate access to the actual meaning of negation, and that this result is not driven by the possibility to fully predict the negated word. However, on the other hand, for the negation that expressed an alternative, in both experiments we found a speed-accuracy trade off, with faster but less accurate answers in true with respect to false conditions. Of course, this speed-accuracy trade-off does not allow us to draw any conclusion about what happened in this case. Indeed, with this type of result we cannot infer whether with the alternative negation participants accessed the affirmative counterpart meaning of the sentence or its actual meaning. As a consequence, this result does not allow us to conclude that describing something that non-exists is a key pragmatic factor that allows to access immediately the actual meaning of the negation. Indeed, since we do not know if with the alternative negation participants accessed the affirmative counterpart or the actual meaning of the negation, we cannot be sure if the occurred access to the actual meaning we found with the nonexistence negation was actually due to the fact that the negation was used to describe something that non-exists and not to other factors, such as for example the possibility to predict the negated word with at least 50% of accuracy and/or the possibility to predict the affirmative meaning of the negative sentences. Moreover, one could propose that another consequence of the speed-accuracy trade off on the alternative negation is that we cannot conclude that the access to the actual meaning of the negation we found with the nonexistence negation is actually an immediate access. Indeed, in our sentence picture verification tasks, we presented the picture 300ms after the sentence. This means that participants performed the truth-judgment task 300ms after reading the sentence. Of course,

we cannot know what happened in those 300ms. However, if, as we predicted, we would have found the polarity by truth-value interaction effect on the alternative negation, this would have ensured that the delay we chose to use between the sentence and the presentation of the image (ISI = 300ms) was not a factor that could explain the access to the actual meaning of the negation and so, in other words, we could, with a certain degree of confidence, exclude that in 300ms participants were able to pass from the affirmative counterpart meaning of the sentence to its actual meaning. With the speed-accuracy trade off, instead, we cannot exclude that also the delay between sentence and picture presentation could be one of the factors responsible for the access to the meaning of the negation. However, in the literature, the prevailing idea is that when interpreted in two stages, the actual meaning of the negation cannot be accessed before 800/1000 ms after the sentence presentation, at least in sentence-picture verification tasks (see for example Luedtke et al., 2008). Thus, in line with this idea, since our ISI was much shorter than 1s, we think that the result we found on the nonexistence negation can be interpreted as evidence of an immediate access to the actual meaning of the sentence.

It is evident that the speed-accuracy trade off had a big impact on the interpretation of our results. For possible, future, follow up experiments it would be important to try to develop paradigms that enable to find effects only on the reaction times variable or only on the accuracy variable. A way to do so could be that to change the presentation duration of the pictures, that in the two experiments we presented here was 250ms. Lengthening the duration until participants' response will increase the possibility to find the effect only on reaction times, reducing the duration of the presentation, for example to 150 ms, will increase the possibility to find the effect only on the accuracies. However, to summarize, despite our results do not provide conclusive evidence that describing something that non-exists is a key pragmatic factor for an immediate access to the actual meaning of the negation, the combined observation of the performance on nonexistence and alternative negation strongly suggests that participants

accessed immediately the actual meaning of the negation that expressed nonexistence and that this access resulted easier than the access to the actual meaning of the negation that expressed an alternative.

Chapter 5: Question Under Discussion[4]

Introduction

Tian, Breheny and Ferguson (2010) proposed that a key pragmatic factor which should allow an immediate access to the negated meaning is the presence of a negative QUD (Question under discussion, Roberts, 1996, 2012) at the discourse level, and implemented a simple manipulation of the syntactic structure of the sentence that, according to their theory, should implicitly elicit a negative QUD in terms of an hypothetic licencing pragmatic context also when the sentence is presented in isolation (out of the blue). With this manipulation Tian et al. (2010) were able to eliminate the advantage for false/mismatching negative sentences, traditionally assumed as a signature of a dual stage of processing. The authors tested their proposal in two experiments. In a first work, Tian et al. (2010), implemented a probe-recognition task, a paradigm that was already used by Kaup et al. (2007) to study negation comprehension, and that, as we have seen in the Introduction Chapter of this thesis, differently from truth-judgment tasks, does not require an explicit truth-value judgment, but simply to evaluate if an object depicted in a picture has been mentioned in a preceding sentence or not. In their experiment, Tian and colleagues (2010), compared the comprehension of two types of negative sentences: negative sentences in the simple SVO (subject verb object) form (e.g. "Jane didn't cook the spaghetti") and negative sentences in the Cleft form (e.g. "It was Jane who didn't cook the spaghetti"). The authors propose that these two types of negative sentences elicit two QUDs which differ in polarity, with negative sentences in the SVO form that elicit a positive QUD and negative sentences in the Cleft form that elicit a negative QUD. In line with the QUD theory by Roberts (2012), a QUD is nothing more than an implicit question that a sentence answers. In this way, the QUD

[4] These studies have been carried out in collaboration with Chiara Finocchiaro and Francesco Vespignani

signals the topic the sentence is talking about and allows to understand which is the discourse in which the sentence is uttered. The QUD theory proposes that when a sentence is presented out of the blue, the listener will automatically try to infer the discourse in which the sentence is uttered and so, in other words, she/he will try to infer which is the QUD to which the sentence answers. Tian et al., (2010) propose that the difference in the polarity of the QUD elicited by SVO and Cleft sentences is due to the fact that Cleft sentences, differently from SVO sentences, are presupposition triggers. For example, a sentence such as "It was Jane who didn't cook the spaghetti" elicits the presupposition that "*someone did not cook the spaghetti*". In line with the QUD theory, Tian and colleagues propose that the presupposition becomes part of the QUD answered by the sentence and so they propose that the QUD for a cleft sentence as "It was Jane who didn't cook the spaghetti" is negative, of the type "Who did not cook the spaghetti?". Differently, as written above, SVO sentences are not presupposition triggers. In light of this Tian et al (2010) propose that, if there are no other contextual factors that signal that this is not the case, it is more common that a sentence answers a positive rather than a negative QUD, and so they propose that a SVO sentence such as "Jane did not cook the spaghetti" answers a positive QUD of the type "*whether Jane cooked the spaghetti*". All the experimental sentences presented by Tian et al. (2010) in their experiment described an object (e.g. spaghetti) in a physical state that can have an opposite (e.g. cooked/uncooked). After each of these sentences, the authors presented a picture depicting the object mentioned in the sentence (e.g. spaghetti). In this picture, the object mentioned could be depicted in the physical state as described by the sentence (e.g. raw spaghetti), match condition, or in the opposite physical state (e.g. cooked spaghetti), mismatch condition. The task of participants was to answer if the object represented in the picture was mentioned in the previous sentence or not, independently from its state, by responding "yes" or "no". It is important to highlight two things about the paradigm used by Tian et al, (2010). First of all, all the experimental sentences presented by the authors were

negative (or in the SVO or in the Cleft form). Secondly, the task of participants was simply to answer if the object was mentioned or not in the sentence, independently from the physical state in which the object was represented in the picture. This means that all the experimental sentences required a "yes" response as they were always followed by a picture in which the object mentioned was depicted. In addition to the experimental sentences, filler sentences that were followed by a picture in which an unmentioned object was depicted were presented. Reaction times showed participants were faster to respond "yes" in the mismatch versus the match condition for negative simple SVO sentences, but they were faster in the match versus the mismatch condition for negative Cleft sentences. The result was interpreted by the authors as evidence that in the case of negative simple SVO sentences, as proposed by the two stage models of negation comprehension, participants accessed the affirmative counterpart meaning of the sentence. On the other hand, in the case of negative Cleft sentences, participants were able to access immediately the actual meaning of the sentence. As anticipated before, the authors attributed the different results between negative sentences in the SVO and in the Cleft form to the different polarity of the QUDs answered by the two sentences. In this way, the authors concluded that if a negative sentence answers a negative QUD, it is possible to access immediately its actual meaning. The validity of this proposal was confirmed in a subsequent work by the same authors (Tian, Ferguson & Breheny, 2016). The authors compared the processing of the same negative sentences used in the first experiment, SVO and Cleft, with their respective affirmative counterpart (e.g. "Jane has cooked the spaghetti" for "Jane hasn't cooked the spaghetti" and "It is Jane who has cooked the spaghetti" for "It is Jane who hasn't cooked the spaghetti"). To do this, they used a visual world paradigm. They presented participants with pictures in which the object mentioned in the sentence was depicted in two opposite physical states. In one part of the screen, the picture corresponded to the physical state described by the sentence, and in another part of the screen the same object was presented in

the opposite physical state. One second after the onset of the pictures a sentence was auditorily presented and participants were instructed to simply listen to it and look at the pictures, the gaze direction was monitored using an eye tracker. In previous studies (Cooper, 1974; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995), it has been shown that even without an explicit task, participants direct their gaze on specific parts of the screen which are relevant with respect to the message conveyed by the spoken utterance. These attention shifts are thought to reflect the changing in the mental representation of events, also in anticipating relevant aspects on the basis of semantic and syntactic properties of the sentence (Altmann and Kamide, 1999, 2007). Tian et al. (2016) compared the looks of participants when presented with negative and affirmative Cleft and SVO sentences. The authors found that with SVO sentences participants directed their looks toward the object depicted in the described physical state faster if the sentence presented was affirmative rather than negative. Indeed, whit affirmative sentences participants immediately directed their gaze toward the object depicted in the physical state described by the sentence. Whit negative sentences, instead, participants initially directed their gaze to both the object depicted in the physical state described and the one depicted in the opposite physical state, without a clear preference, and only later they directed their looks toward the object depicted in the physical state described by the sentence. The pattern was again different for Cleft sentences. In this case, participants were equally fast with affirmative and negative sentences to direct their gaze toward the object depicted in the described physical state. The result with SVO negative sentences was interpreted as in line with the proposal of the two-stage models of negation comprehension. Indeed, SVO negative sentences showed to require extra-processing with respect to affirmative SVO sentences, and the authors proposed that this extra processing was due to the initial activation of the affirmative counterpart meaning of the negation. On the other hand, negative Cleft sentences showed to not require any extra processing with respect to affirmative Cleft sentences, and so

the result seems to indicate that in this case the actual meaning of the negative sentences was immediately accessed. Tian et al. (2010, 2016) results are impressive in showing how the use of syntactic structures which convey specific pragmatic function allow the processing of negation to drastically change, without manipulating world knowledge dimensions, differently from Nieuwland and Kuperberg (2008) where pragmatic licensing was only empirically defined. Despite the theoretical relevance of these results, no other study to our knowledge was conducted using this manipulation and no direct replication of this research by an independent research group is published. In the first experiment of this chapter, we adapted the Cleft manipulation to the sentence-picture paradigm developed in the first part of this thesis, in the second experiment we implemented a more direct replica of the paradigm by Tian et al. (2010).

Experiment 4

In the first experiment we implemented the syntactic manipulation by Tian, Breheny and Ferguson (2010) on a sentence-picture verification task using coloured shapes as in the previous experiments (Chapters 3;4). As in Tian et al. (2010), the sentences presented might be in SVO or in Cleft form. After each sentence we presented a picture that might represent correctly (true conditions) or not (false conditions) the meaning of the sentence. The task of participants was to respond "true" if the sentence correctly described the picture and "false" otherwise. In line with the proposal by Tian et al., (2010) we predicted an access to the meaning of the affirmative counterpart for SVO negative sentences (two-stages) but evidence for a direct access to the actual meaning of the sentence for Cleft negative sentences (one-stage). Within our paradigm this hypothesis implies faster and/or more accurate responses in false with respect to true conditions for negative SVO sentences, and faster and/or more accurate responses for true with respect to false conditions for negative Cleft sentences. Initially this experiment was thought as a pilot experiment for an EEG study aimed to measure the N2pc component elicited

by the picture presentation, given this component is elicited by lateralized presentation of the targets the coloured shapes will not be intermixed on the screen as in previous studies but objects which are homogeneous for shapes and colour will be confined to one side of the screen. A design like that could also allow to adapt the paradigm to a visual word eye-tracking experiment, similar to Tian et al. (2016). Due to Covid-19 situation, neither EEG or eye-tracking experiments were run.

Methods

Participants

We recruited 12 participants. The data collection of this pilot experiment was stopped by the lockdown situation due to the Covid-19 pandemic in Italy. At the same time, we think 12 can be considered a sufficient numerosity based on Calignano (2020). In the experiment reported there, the authors studied the same ERP component we would have studied in our EEG experiment and, by using the same paradigm and similar images stimuli to those used in our experiment, found the statistical three-way interaction we were looking for, including 12 participants. Participants were all native speakers of Italian, with normal or corrected to normal vision, without any neurological or psychological disease, and within an age range between 18 and 40 years old. If they were students of the Department of Psychology and Cognitive Science of the University of Trento, participants received course credits for their participation.

Stimuli

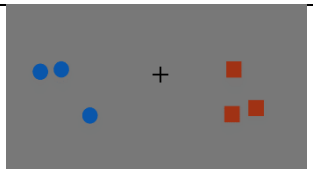
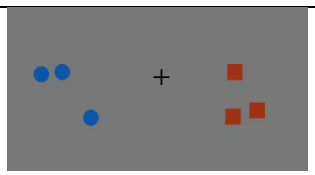
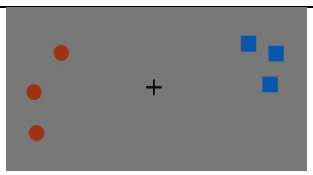
The experiment was a sentence – picture verification task. Thirty-two experimental sentences, 16 in the simple SVO form (8 affirmatives and 8 negatives) and 16 in the Cleft form (8 affirmatives and 8 negatives), were constructed. All the sentences were in Italian. The SVO

sentences were of the form “I X sono/non sono Y” (“The X are/are not Y”). The Cleft sentences were of the form “Sono i X che sono/non sono Y" (“They are X that are/are not Y”). The noun in subject position (X) could be “cerchi” (“circles”) or “quadrati” (“squares”). In the adjective position (Y) there could be one of four colour terms (“verde”, “rosso”, “blu” or “viola”/ “green”, “red”, “blue” or “violet”).

The visual display was created using a grey (RGB: 120,120,120; 25 cd/m2) background, a black fixation cross (32 X 32 pixel; visual angle 0° 36') at the centre of the picture and two groups of geometrical objects with different shapes and colour, each type segregated on one side of the display (see Table 1 for a display example). In half of the displays the circles were on the left, in the other half the squares were on the left. Each group of shapes could be coloured with one of four different colours: green (RGB:16,102,24), red (RGB:160,50,20), blue (RGB:9,86,171) or violet (RGB:123,54,142). Circles had a diameter, and squares a side, of 32 pixels (visual angle 0° 36'). In total we created 1.512 pictures divided in 36 tokens per every of the 32 possible combinations of shapes (circle vs squares)/position (left vs right of the fixation cross)/colours (green vs red vs blue vs violet). The images were constructed using two 3x3 (visual angle 4°32'X2°50') matrixes. The two matrixes were displaced 125 pixels (visual angle = 2° 22') away from the centre of the fixation cross, one on the right side and other one on the left side of the cross. Within a matrix, the shapes could occupy randomly three of the nine cells. In a cell, the centre of the shape was then randomly jittered away from the centre of the cell along both horizontal and vertical axes, using a uniform probability distribution of jitter amplitude between 0 and 0.8 of half the axis length.

Design

Combining each sentence (affirmative or negative, SVO or Cleft) with a picture in which at least one of the two groups of shapes was coloured with the mentioned colour, we create true and false conditions for both affirmative and negative sentences as in Table 1. The resulting experimental design is thus a 2 (Polarity: Affirmative vs Negative) by 2 (Sentence_Type: SVO vs Cleft) by 2 (Truth_Value: True or False) within-subjects design. The experiment was divided in 8 blocks of 36 trials each. Within a block only one of the two levels of the Sentence Type factor (SVO or Cleft) was presented, while all the other experimental factors were randomly presented. It is important to outline that among the total 288 trials of the experiment, 32 trials were catch trials in which the two groups of shapes had the same (mentioned) colour. These catch trials were one for each of the 32 possible combinations of Sentence_Type, Truth_Value, Subject of the sentence (circles or squares), Position of the subject of the sentence in the image (left vs right side of the fixation cross). These catch trials have been constructed to avoid participants attended only one side of the screen to correctly respond to all trials. In 144 out of 288 trials, the subject of the sentence were circles, in other 144 were squares. Of these 144, in 72 of them the subject of the sentence was depicted on the left of the fixation cross, in the other 72 on the right.

SENTENCE	IMAGE True Condition	IMAGE False Condition
I cerchi sono rossi *The circles are red*		
I cerchi non sono rossi *The circles are not red*		

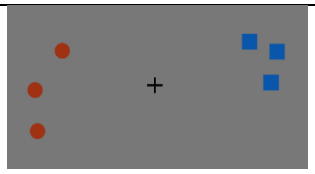
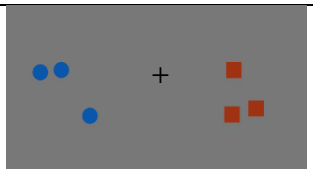
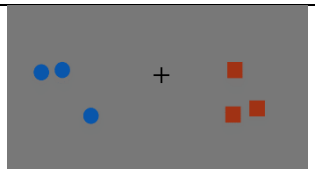
Sono i cerchi che sono rossi *They are the circles that are red*		
Sono i cerchi che non sono rossi *They are the circles that are not red*		

Table 1. Example of the experimental conditions within the experiment.

Procedure

Participants were sited in a quiet room with the luminance of the room kept at 5 cd/m2. The computer screen was placed 80 cm in front of the participants. To reduce head movements, subjects were asked to use a chin-rest. The materials were displayed using the PsycoPy3 Software (Peirce, Gray, Simpson, MacAskill, Höchenberger, Sogo, & Lindeløv 2019).

Each trial began with a blank grey monitor for 1000 ms followed by a black fixation cross on a grey background for 500 ms. Then, immediately below the fixation cross, the sentence was presented word by word, with each word presented for 300 ms followed by a screen with only the fixation cross for 100 ms. After the last word of the sentence the fixation cross remained on the screen for 1000 ms, after which the picture was presented (SOA=1400ms after the onset of the last word) for 200 ms. After the offset of the image participants were required to respond. If they didn't within 4s a message "Too slow. Try to answer faster" appeared and the subsequent trial started.

Before starting the experiment, participants were assigned to one of two different groups: one group in which the first four blocks contained only sentences in the SVO form and the last four blocks only sentences in the Cleft form, or one group in which the first four blocks contained only sentences in the Cleft form and the last four only sentences in the simple SVO form. Six participants were assigned to the first group and six to the second one. In addition, participants

were instructed to keep their gaze on the fixation cross both during the sentence and the picture presentation and to indicate as quickly and accurately as possible whether the picture correctly represented or not the sentence meaning, by pressing one of two buttons, labelled as "true" and "false". Half of the participants made "true" responses with their left index finger (pressing the “z” key on the pc keyboard) and "false" responses with their right index finger (pressing the “m” key). For the remaining participants response hands were reversed. Each test session started with a practice block of 12 trials, 6 with cleft sentences and 6 with simple sentences, randomly presented. During the practice block participants received feedbacks on their response, whereas no feedback was given on response accuracy during experimental blocks.

Analysis and Results

We analysed reaction times (RTs) and accuracy in the truth-judgment task using the lme4 package (Bates, Mächler, Bolker & Walker, 2015) in Rstudio software, 1.2.1335 version (RStudio Team, 2018).

Reaction Times

Reaction times of wrong responses were excluded from the analysis. On the remaining data were fitted 8 different Generalized Linear Mixed Effect Models. All the models had an inverse gaussian distribution (Lo and Andrews, 2015) and a random effects structure composed by random intercepts for participants. (see Table 2).

Models	**AIC**	**R Df**	**R Dv**	**Δ AIC**	**AICw**	**Df**
RTs ~ (1|Participants)	2808.4	2810	1085.	649.	0	3
RTs ~ Polarity+Truth_Value + (1| Participants)	2247.6	2808	887.	87.9	0	5
RTs ~ Polarity+Truth_Value+Sentece_Type + (1| Participants)	2240.7	2807	884.	81	0	6

RTs ~ Polarity*Truth_Value + (1| Participants)	2166.8	2807	861.	7.1	0.02	6
RTs ~ Polarity+Truth_Value*Sentence_Type+ (1| Participants)	2242.3	2806	884.	82.7	0	7
RTs ~ Polarity*Truth_Value+Sentence_Type + (1| Participants)	2159.7	2806	858.	0	0.81	7
RTs ~ Truth_Value + Polarity*Sentence_Type + (1| Participants)	2240.6	2806	883.	80.9	0	7
RTs ~ Polarity*Truth_Value*Sentence_Type + (1| Participants)	2162.8	2803	857.	3.1	0.17	10

Table 2. Models' comparison for Reaction Times. AIC = Akaike Index, R Df = Residual Degree of Freedom, R Dv = Residual Deviance, Δ AIC = delta Akaike, AICw = Akaike weight, Df = Degree of Freedom.

The models' comparison among Generalized Linear Mixed Effect Models showed that the model with the main effect of Sentence_Type and the interaction between Polarity and Truth_Value was the model which better approximate the data (lower AIC index). To better describe these effects, we conducted paired comparisons on the model selected, p-values were adjusted with Tuckey correction for multiple comparisons. Paired comparisons revealed that with affirmative sentences participants were faster to respond in true with respect to false condition (β = -0.2114, SE = 0.0153, z.ratio = -13.781, p <0.0001), while with negative sentences they were faster in false with respect to true condition (β = -0.0816, SE = 0.0279, z.ratio = -2.923, p = 0.0069). Lastly, the model predicted that participants were always faster with SVO sentences than with Cleft sentences (β = - 0.0333, SE = 0.0111, z.ratio = - 3.001, p = 0.0027) (see Figure 1).

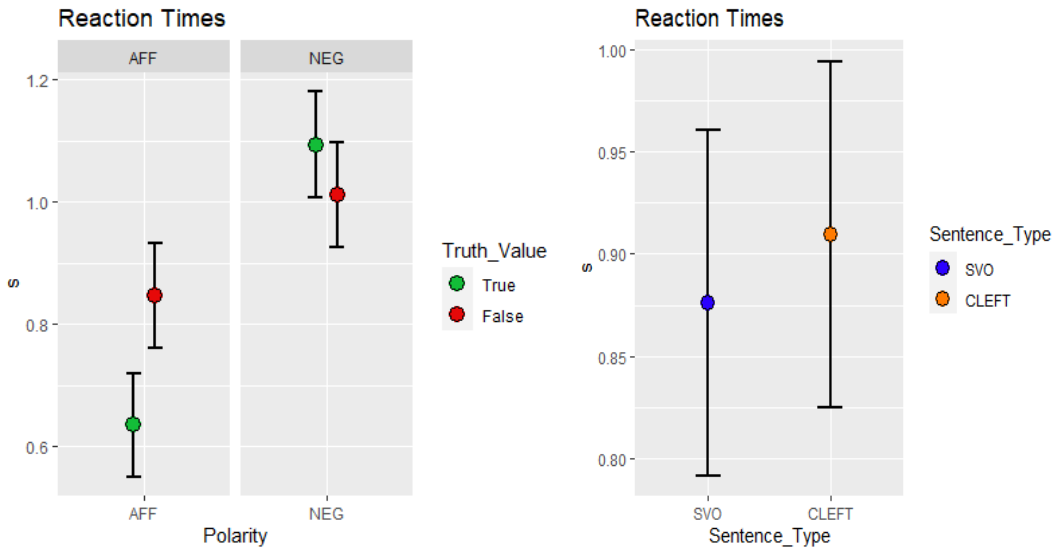


*Figure 1. Plots of the predicted Reaction Times by the model Polarity*Truth_Value+Sentence_Type + (1| Participants).*

Accuracy

On the responses we fitted eight Generalized Linear Mixed Effect Models. The distribution used was the binomial distribution and the random effects' structure was composed by random intercepts for participants (see Table 3)

Models	**AIC**	**R Df**	**R Dv**	**Δ AIC**	**AICw**	**Df**
ACC ~ (1| Participants)	1675.8	3057	1640.	42.2	0	2
ACC ~ Polarity + Truth_Value + (1| Participants)	1657.1	3055	1617.	23.6	0	4
ACC ~ Polarity + Sentence_Type + Truth_Value + (1| Participants)	1647.6	3054	1606.	14.1	0	5
ACC ~ Polarity * Truth_Value + (1| Participants)	1643.1	3054	1601.	9.6	0.01	5
ACC ~ Polarity + Truth_Value * Sentence_Type + (1| Participants)	1647.5	3053	1604.	14	0	6
ACC ~ Polarity * Truth_Value + Sentence_Type + (1| Participants)	1633.5	3053	1590.	0	0.68	6
ACC ~ Truth_Value + Polarity*Sentence_Type + (1| Participants)	1648.5	3053	1605.	15	0	6
ACC ~ Polarity * Truth_Value * Sentence_Type + (1| Participants)	1635.1	3050	1585.	1.6	0.31	9

Table 3. Models' comparison for Accuracy. AIC = Akaike Index, R Df = Residual Degree of Freedom, R Dv = Residual Deviance, Δ AIC = delta Akaike, AICw = Akaike weight, Df = Degree of Freedom.

The models' comparison among Generalized Linear Mixed Effect Models showed that, also for accuracy, the model with the main effect of Sentence_Type and the interaction between Polarity and Truth_Value was the model which better approximated the data (lower AIC index). To better described these effects, we conducted paired comparisons on the model selected, p-values were adjusted with Tuckey correction for multiple comparisons. Paired comparisons revealed that with affirmative sentences participants were more accurate to respond in true condition with respect to false one (β = 0.660, SE = 0.230, z.ratio = 2.873, p = 0.0081) while with negative sentences they were more accurate to respond in false condition with respect to true one (β = 0.479, SE = 0.173, z.ratio = 2.773, p = 0.0111). Lastly, for the main effect of the factor Sentence_Type, the model predicted higher accuracy for SVO than for Cleft sentences (β = 0.465, SE = 0.137, z.ratio = 3.390, p = 0.0007) (Figure 2).

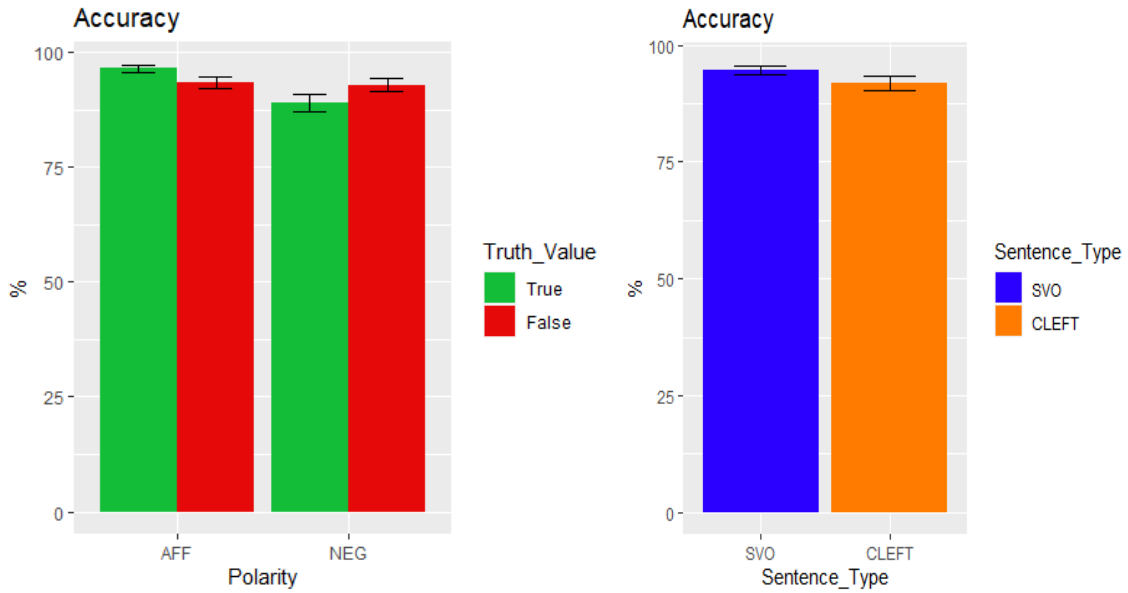


*Figure 2. Plots (right) of the predicted Accuracies for the interaction between Polarity and Truth_Value and (left) for the main effect of Sentence_Type by the model Polarity * Truth_Value + Sentence_Type + (1| Participants)*

Discussion

The aim of experiment 4 was to try to replicate the findings by Tian, Breheny and Ferguson (2010) by using a sentence-picture verification task. To do this we implemented an experiment in which we presented affirmative and negative sentences in the SVO and in the Cleft form. The sentences were about the colour of some geometrical shapes. Each sentence was paired with a picture that can represent correctly (true conditions) or not (false conditions) the sentence meaning. With respect to negative sentences, in line with the proposal and the results reported by Tian et al. (2010), we predicted to find evidence for an access to the affirmative counterpart meaning of the negative sentences presented in the SVO form, and for an access to the actual meaning of the negative sentences presented in the Cleft form. This, within our experimental paradigm, would have translated in faster and/or more accurate responses in false conditions with respect to true ones for SVO sentences and faster and/or more accurate responses in true conditions with respect to false ones for Cleft sentences. However, contrary to our predictions, our models predicted faster reaction times and higher accuracy in false conditions with respect to true ones for both types of negative sentences, independently of the form, SVO or Cleft. Thus, the result seems to indicate that in our experiment, independently of the syntactic form of the sentence, participants accessed the affirmative counterpart meaning of the negative sentences, in line with the classical two-stage models of negation comprehension. Additionally, our analysis shows faster reaction times and higher accuracy for all the sentences in the SVO form with respect to the Cleft form. This seems to indicate that participants were facilitated in our task when the sentence was in the SVO rather than in the Cleft form. This result can be due both to the larger number of words in the Cleft condition (2 extra function words) and its syntactic complexity (an extra subordinate relative clause), but also to the more complex situational representation the reader can build on the basis of a Cleft sentence. Indeed, a Cleft sentence like “Sono I cerchi che sono rossi” (“They are the circles

that are red") triggers the implicature that other shapes different from circles have a different colour than red. This type of implicature is not triggered by a sentence like "I cerchi sono rossi" ("The circles are red"). Given the rather long SOA between the end of sentence presentation and the onset of the picture (1400 ms) we are more prone to attribute this effect to the semantic and pragmatic consequence of the Cleft on the situational representation than to differences in surface complexity since we think that 2 extra function words and a more complex structure without unpreferred attachments, it is unlikely to spill-over to a subsequent judgement task on a picture.

Experiment 5

In Experiment 4 we did not find evidence that Cleft sentence systematically allows for an immediate access to the negative meaning of a sentence as it could have been predicted on the basis of Tian et al. (2010, 2016). Indeed, the finding that for both SVO and Cleft negative sentences, true condition shows a cost with respect to false can be interpreted as evidence that the meaning of affirmative counterpart of the negative sentence is somewhat accessed. Clearly the absence of a 3-way interaction can be due to the low power in the paradigm or to the fact that the negative QUD would be harder to be automatically elicited with a paradigm dealing with massive repetition of same lexical items referring to a rather abstract and impoverished semantic world. In our experiment, in fact, we always dealt with 2 shapes and 4 colours, and this may confound possible QUD across trials. Differently, Tian et al (2010, 2016) worked with more naturalistic situations (e.g. cooking spaghetti) which varied from trial to trial allowing the reader to imagine independent worlds for each sentence that could induce rather independent QUD by pragmatically projecting a possible world in which that sentence could make sense. Further difference between our Experiment 4 and Tian et al. (2010, 2016) are: explicit truth judgment task (our experiment) versus implicit task (picture monitoring in Tian

et al., 2010; visual world paradigm in Tian et al., 2016); the use of adjectives without opposite (colour, our experiment) versus predicates describing a physical state that can have an opposite (cooked/uncooked; open/close, Tian et al. 2010, 2016); the blocked presentation of Cleft and SVO (our experiment) versus the intermixed presentation of the two types of sentences (Tian et al. 2010, 2016). Before trying to tear apart which among these possible differences between Experiment 4 and Tian et al. (2010, 2016) make clefting more or less effective in eliciting a QUD which allows immediate access to negative meaning of a sentence we decided to try to replicate the original study by Tian, Breheny and Ferguson (2010) in Italian. After inspecting the material of the original experiment[5] we realized that some items should be changed since the translation in Italian could have been problematic (e.g. use of compound words) moreover we observed that the pictures were coloured pictures. Given that balance of the physical properties of pictorial material is rather complex and could introduce extra noise in the data we decided to create a new set of stimuli, inspired to the original ones but substituting some sentences and by using line drawing which will be specifically created for the experiment. This latter choice was done in order to minimize differences in the processing needed to parse the pictures. We think that controlled line drawing could have been more effective in testing the reliability of the original experiment and, moreover, could be more easily adapted to other paradigms, using EEG or eye-tracking (visual world) for which physical differences in the pictorial material could be even more relevant. We think this choice has been especially useful since the experiment ended to be run online because of the covid pandemic and differences in the luminosity and dimensions of the screens used by participants may introduce further irrelevant variance in the data when using complex-coloured pictures than line drawings.

[5] We want here to acknowledge prof. Richard Breheny and dr. Yen Tian who shared with us the material of their original experiment, including both the sentences and the pictures in both physical states used in their work Tian Breheney and Ferguson (2010).

Methods

Participants

We recruited 200 participants for this experiment. The sample size was exploratory. Participants were recruited via Social Media (Facebook and Instagram) through posts on unofficial pages of Italian Psychology Universities (Trento, Bergamo, Padova) and on pages which advertise psychological experiments. These posts asked participants to contact the experimenter by sending him an e-mail contact in the case they were interested to participate in a psycholinguistic on-line experiment. Advertisement included information about experiment duration (25 minutes) and reward for participation (3.5€) and underlined the following inclusion criteria: age between 18-40, Italian first language speakers, no reading disorders, normal or correct to normal vision, no neurological disorders, not to currently assume drugs with side effects on the psychological functioning.

Stimuli

The experiment was a probe-recognition task. As in Tian et al. (2010), 28 experimental items were constructed. All the experimental items were composed in Italian, some of which as a direct translation of the original items and contained a negation ("non"/ "not"), a verbal predicate that described a physical state that can have an opposite (e.g. "allacciare"/"slacciare"; "fasten"/"unfasten") and one object (e.g. "scarpe"/ "shoes"). For each of the 28 experimental items two different syntactic forms were constructed: a simple SVO form (e.g. "Luca non ha allacciato le scarpe"/"Luca didn't fasten the shoes") and a Cleft form (e.g. "É Luca che non ha allacciato le scarpe"/"It is Luca who didn't fasten the shoes"). For each of the 28 experimental items 2 images were constructed. Both images represented the object mentioned in the sentence, one of these represented it in the physical state described by the sentence (e.g. untied shoes) and the other one representing it in the opposite physical state (e.g. tied shoes). In

addition to these experimental stimuli, other 56 filler sentences and 56 filler images were constructed. The 56 filler sentences were composed by 14 negative sentences and 42 affirmative sentences, in order to have the same number of affirmative and negative sentences in the whole experiment. Both for affirmative and negative sentences, half of the sentences were in simple SVO form and half in the Cleft form. Differently from the experimental sentences, and as in the original study by Tian et al. (2010), the filler sentences did not always contain a verbal predicate that described a physical state that can have an opposite, but always contained an object (e.g. "Eva ha comprato il furgoncino"/ "Eva bought the van"). The 56 filler images were composed by 28 images that depicted the object mentioned in 28 (out of 42) affirmative filler sentences, while the other 28 pictures depicted an unmentioned object for the 14 filler negative sentences and for the remaining 14 filler affirmative sentences. Since the participants' task was to respond "yes" if in the image was represented the object mentioned in the sentence (independently from its physical state) and "no" if was represented an unmentioned object, the use of these filler images ensured that the "yes" and "no" responses in whole experiment were balanced. Given that the task by itself (probe recognition) does not require a careful reading and understanding of the sentence and the possible probe was always the final word, as in Tian et al (2010), some comprehension question was randomly added in order to boost attention to sentence meaning. For 28 sentences (9 experimental items and 19 filler sentences), a yes/no comprehension question was constructed in a way that a full comprehension of the sentence meaning was necessary to respond correctly (for example, for a sentence such us "É Leo che non ha inzuppato la brioche" ("It is Leo who did not soak the brioche") the question was "La brioche di Leo è ancora asciutta?" (Is the brioche of Leo still dry?")). In this way, the questions ensured that participants read and understood all the sentence and they did not focus only on the name of the object mentioned.

As written above, the more relevant difference between our paradigm and stimuli set and those implemented by Tian et al. (2010) is the use of different pictorial material. Indeed, in order to minimize irrelevant variance and to avoid possible confounds on reaction times due to systematic uncontrolled differences in the visual complexity of the images, we decided to construct black and white line drawing instead of coloured pictures, these pictures were drawn from scratch by a professional web illustrator and reviewed by our research group, trying to keep similar levels of details and complexity across items.

Two additional minor differences between our experiment and the original one regarded the sentences. Indeed, differently from the original study, within the 28 experimental items we decided to never repeat a verbal predicate, and so to use 28 different verbs. Moreover, for the sentences in the Cleft form we decided to use the present tense, "It is X who …", instead of the past tense, "It was X who…", used by Tian et al. (2010). This decision was taken since in Italian the most common version of the Cleft sentence at the past is "É stato X a …"/ "It was X to …" rather than a relative clause "É stato X che …"/ "It was X who …". Among all different changes in the structure, we decided that turning the Cleft to present could be more similar to the original English Cleft sentences used by Tian et al. (2010).

The whole stimulus set is available upon email request to the author.

Design

Both our two experimental factors, items and pictures, had two levels. Items could be in the SVO or in the Cleft forms and pictures could represent the mentioned object either in the correct or in the opposite physical state (match/mismatch condition). The 28 experimental items were distributed within this 2 (Sentence_Type: SVO vs Cleft) x 2 (images: match or mismatch) in a within-subjects experimental design (see Table 4) using a latin square design in order to avoid item repetition in different cells within participants and allowing an across subject item

balance. The 200 participants have been equally distributed between the four presentation lists (50 participants per list). Within each list the sentences were presented in random order.

SENTENCE	IMAGE Match Condition	IMAGE Mismatch Condition
Luca non ha allacciato le scarpe Luca didn't fasten the shoes		
É Luca che non ha allacciato le scarpe It is Luca who didn't fasten the shoes		

Table 4. Example of the experimental conditions in the experiment.

Procedure

Once a participant accepted to take part in the experiment an email with the link of the experiment and all the instructions to set up the environment to perform the experiment was sent to him/her. The instructions specified that the experiment was to be done in in a quiet room in which participants could sit in front of a computer (No Tablet, No Smartphone) laid on a table. Moreover, we asked them to switch off all the devices and computer apps that could distract them, to keep the browser page full screen, to keep open only the page on which the experiment would be run, and to make sure that on the computer monitor there weren't light reflections that could make the reading difficult. At the end of the e-mail participants found the experiment link. Clicking on that link participants were redirected to the Jatos software (Lange, Kühn, & Filevich, 2015) on which the experiment run. The material of the experiment was

displayed using the OpenSesame On-line Software (Mathôt, Schreij, & Theeuwes 2012). At this point, the instructions about the experimental task were displayed on screen. The instructions explained that the participants' task was to read a sentence and, as soon as they finished to read it, to press the space bar so that a picture appeared. They were informed that once the picture appeared, their task was to respond as faster and accurately as possible "yes", by pressing the "o" key, if in the picture the object mentioned in the sentence was represented, or "no", by pressing the "p" key, if in the picture a different object to that mentioned in the sentence was represented. It is important to outline participants were instructed to answer "yes" ("o" key) if the object represented was mentioned in the previous sentence, independently from the physical state in which the object was depicted in the image. Moreover, they were informed that sometimes and in an unpredictable manner, after their response to the image, a question about the meaning of the sentence they just read could appear. They were instructed that answer keys to the questions were the same of the previous task ("yes" = "o" key/"no" = "p" key) and that they must respond as accurately as possible but that it was not necessary to be fast. After reading the instructions participants could start the experiment. Each trial of the experiment started with 500ms of blank screen followed by a fixation cross presented in the centre of the screen for 1200ms. Subsequently the sentence was presented in the centre of the screen. The sentence lasted on the screen until participants clicked the space bar on their computer keyboard. Once participants clicked the space bar, a fixation cross appeared for 250ms and then a picture (approx. 7.5X7.5 cm) was presented in the centre of the screen. At this point, the picture lasted on the screen until the participant's response. The comprehension question, when presented, appeared after 500ms of blank screen and remained on the screen until the participant's response. After the end of a trial the following trial started automatically. To familiarize with the task, before starting the real experiment participants had to complete a practice block composed by 6 trials (2 with comprehension questions). During this practice

block, after each response, participants received feedback about their accuracy. After the practice block was completed, participants could decide whether to start the experiment or, if something was not clear, to repeat the practice block a second, and only a second, time. Differently from the practice block, during the experiment participants did not receive any feedback on their responses. The experiment took about 25 minutes to complete.

Analysis and Results

Since in the original study by Tian et al. (2010) no participant answered to the comprehension questions with an overall accuracy lower than 80% (see Tian & Breheny (2010)), we decided to exclude participants who performed under this threshold from our analysis, leading to excluding 27 participants out of 200. As in Tian, Breheny and Ferguson (2010) we analysed only reaction times (RTs) in the task for the experimental items. We analysed them using the lme4 package (Bates Mächler Bolker & Walker 2015) in Rstudio software, 1.2.1335 version (RStudio Team, 2018).

Reaction Times

Reaction times slower than 3000ms and faster than 300ms (cf. Tian et al., 2010) were excluded, as well as wrong (no) responses. The remaining reaction times so corrected we fitted five different Generalized Linear Mixed Models with an inverse gaussian distribution (Lo & Andrews, 2015) and a random structure composed by random intercept for Participants and Items (see Table 5). We performed a model selection between these models fitted using AIC index as decision criterion.

Models	AIC	R Df	R Dv	Δ AIC	AICw	Df
RTs ~ (1|Participants) + (1|Items)	65867	4759	0.637	0	0.34	4
RTs ~ Sentence_Type + (1| Participants) + (1|Items)	65867	4758	0.636	0.2	0.31	5
RTs ~ Match_Mismatch + (1| Participants) + (1|Items)	65869	4758	0.637	2	0.13	5
RTs ~ Sentence_Type + Match_Mismatch + (1| Participants) + (1|Items)	65869	4757	0.636	2.2	0.11	6
RTs ~ Sentence_Type * Match_Mismatch + (1| Participants) + (1|Items)	65870	4756	0.636	2.3	0.11	7

Table 5. Models' comparison for Reaction Times. AIC = Akaike Index, R Df = Residual Degree of Freedom, R Dv = Residual Deviance, Δ AIC = delta Akaike, AICw = Akaike weight, Df = Degree of Freedom.

Two models presented the lowest AIC index (AIC = 65867). These two models were the Null model and the model with the main effect of Sentence_Type. Since the AIC index was the same, but the Null model is simpler than the model with Sentence_Type as main effect, we selected the first one. The selection of the Null model implies that no one of the experimental factors presented affected the speediness to which participants performed in the task (see Figure 3).

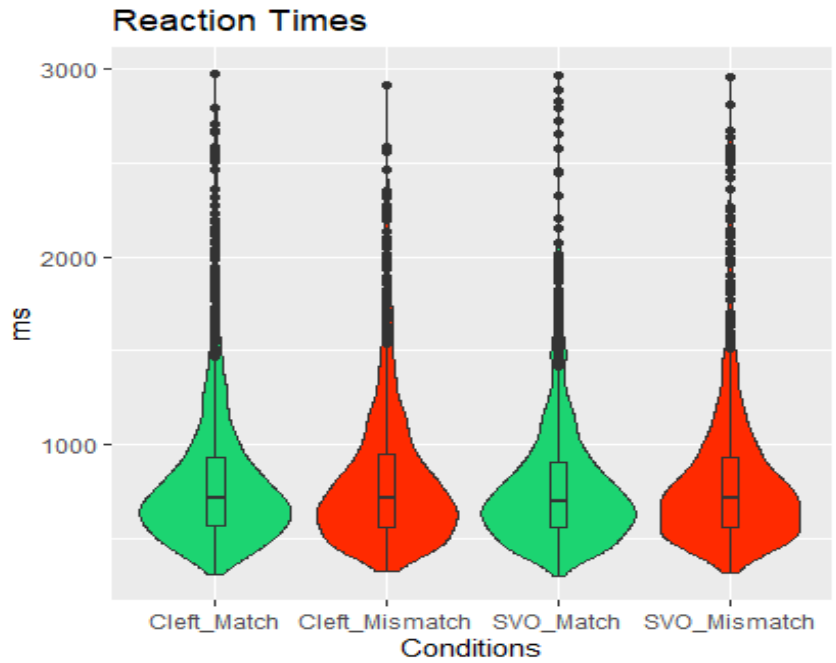


Figure 3. Violinplots of the mean Reaction Times in the different experimental conditions.

Discussion

In this second experiment we run a replica in Italian of the experiment by Tian et al. (2010). We implemented a probe-recognition task in which participants had to decide whether a depicted object was mentioned in a negative SVO or Cleft sentence independently of its physical state: same (match condition) or different (mismatch condition). We predicted to replicate the results by Tian et al (2010), that is faster RTs in the mismatch versus the match conditions for SVO negative sentences, and faster RTs in the match versus the mismatch conditions for Cleft negative sentences. However, contrary to our predictions, our results indicated that for both SVO and Cleft negative sentences participants were equally fast in the match and in the mismatch conditions. In other words, we did not find any difference between the experimental conditions. Of course, this result cannot be taken as a successful replication of the original study by Tian et al., (2010). Several can be the reasons for this failure. For example, differently from the original study, we presented sentences in Italian and our Cleft sentences were conjugated at the present tense, instead that at the past tense as in the original study. We think, however, that these two reasons cannot explain our failure replication. First of all, to our knowledge, the are no reasons for which in Italian the effect on Cleft sentences should disappear. Indeed, also in Italian, Cleft sentences are presupposition triggers and so, in line with Tian and colleagues' proposal, should elicit a negative QUD. Moreover, also the conjugation of the verb should not be a problem since, if it is true that in Tian et al. (2010) the authors used Cleft sentences at the past tense, it is at the same time true that in Tian et al., (2016) they used Cleft sentences at the present tense, and they found effects in line with their proposal. More importantly, we think the Italian language and the verb conjunction are not valid reasons to explain our replication failure because we did not simply do not replicate the results by Tian et al. (2010) with Cleft negative sentences, but we did not find any effects due to experimental factors. This means that we did not even find, neither with Cleft sentences nor

with SVO sentences, the classical effect that is usually found in probe-recognition tasks with negative sentences, and so that participants are faster to respond in the mismatch vs the match condition. In the light of the lack of this effect (at least with SVO sentences) we think that the reason for our failure is to search in the other two biggest differences between our paradigm and that adopted by Tian et al. (2010) that are: the use of black and white line drawing pictures instead of coloured ones and the fact that our experiment was run online. We think that both the reasons (or the combination of the two) can be a valid explanation for our replication failure. We suggest, indeed, that the coloured pictures used by Tian et al. (2010), having a greater visual complexity with respect to the black and white pictures we used, required more time to be parsed. This additional parsing time gave participants more time to interpret the sentences and so, maybe, allowed the emergence of the effect. This idea seems to be supported by the higher mean reaction times found by Tian and colleagues in their experiment (1000ms) vs the mean reaction times found in our replica (900ms). On the other hand, the on-line administration can justify our failure because the effects found with the probe-recognition task are usually small so, maybe, the "noisiness" of the on-line measurement could have prevented the emergence of the effect. Moreover, a paper by Zwann and Pecher (2012) has shown that it is not easy to replicate on-line the results found in lab with this paradigm. Despite, as we have written above, we think that both the reasons can be valid explanations for our replication failure, since our paradigm is not the first probe-recognition task which used black and white pictures to investigate negation comprehension, and that these other experiments found the classical effect of faster answers in the mismatch vs the match condition with negative sentences (e.g. Kaup, Yaxley, Madden, Zwaan, & Lüdtke, 2007), if we have to bet on one of the two explanations we think that the main issue in our experiment could have been its on-line administration.

General Discussion

In this chapter we tested in two experiments the proposal by Tian et al. (2010) according to which if a negative sentence answers to a negative Question Under Discussion (QUD) it is possible to immediately access its actual meaning. To test this proposal, we used the same experimental manipulation used by Tian et al., (2010), and compared negative sentences in simple SVO (e.g. "Jane didn't cook the spaghetti") and in Cleft ("It was Jane who didn't cook the spaghetti") forms. Tian and colleagues suggested that negative sentences in the SVO form answer a positive QUD (e.g. "whether Jane cook the spaghetti") while negative sentences in the Cleft form answer a negative QUD (e.g. "Who didn't cook the spaghetti?"). In experiment 4 we used a sentence – picture verification task, while in experiment 5 we run a faithful replication in Italian of the original study by Tian et al. (2010) by using a probe-recognition task. In both experiments we failed to find positive evidence for a direct access to the actual meaning of negative sentences in the Cleft form, as assumed by Tian et al. (2010). Indeed, in experiment 4, both for negative sentences in the SVO and in the Cleft form we found evidence for an access to the affirmative counterpart meaning of the sentences. This result indicates that, within our experiment, the affirmative counterpart meaning of the sentence was accessed independently from the polarity of the QUD answered by the negative sentences or that the paradigm was not strongly inducing a negative QUD. In experiment 5, instead, not only we did not find evidence for an access to the actual meaning of negative Cleft sentences but neither of the access to the affirmative counterpart meaning, both for SVO and Cleft negative sentences. Given the lack of this classical and solid effect of the access to the affirmative counterpart meaning of the sentence (at least in the SVO form) and given that the major difference between our experiment and the others that have found this result within this paradigm (Kaup, Luedtke, Zwaan, 2005; Kaup, Yaxley, Madden, Zwaan, & Lüdtke, 2007; Tian et al., 2010) was that our experiment was run on-line, we think that probably our result should be interpreted as evidence

that the paradigm used by Tian et al., (2010) is not solid enough to study negation in on-line. In light of this we think that new, more solid and feasible paradigms should be developed to investigate the Tian and colleagues' proposal.

To conclude, we think that, overall, our results cannot be taken as a falsification of the theoretical proposal of Tian et al. (2010) but as evidence that the effects found by the author (2010, 2016) are not easy to replicate, at least in on-line experiments that use black and white line drawing images, and generalize, at least to experimental paradigms that share the differences we introduced in our sentence-picture verification task (experiment 4). In particular we hypothesize that the most significant differences between our experiment 4 and the experiments by Tian and colleagues (2010, 2016) could be the presence in our experiment of an abstract and impoverished semantic world, of an explicit truth-judgment task and the use of unary predicates (e.g. "not red" in our experiment could mean green, blue or violet while Tian et al. (2010, 2016), in their experiments, used verbal predicates that have an opposite such us open/close; switch off/switch on, etc.). Anyway, we think it is worth to continue to investigate and extend the theoretical proposal by Tian and colleagues, also trying to identify if characteristics of the Question Under Discussion other than its polarity can affect the interpretation of negation. For example we think that the idea of Tian et al., (2010, 2016) to compare negative sentences in the SVO form (e.g. "Jane did not cook the spaghetti") with negative sentences in the Cleft form (e.g. "It was Jane who didn't cook the spaghetti") in order to implement a comparison with negative sentences that answer a positive and a negative QUD could have had the consequence to manipulate also another characteristic of the QUD. In particular, we think that for negative SVO sentences there can be an ambiguity on which is the type of QUD answered by the sentence that is not present for negative Cleft sentences. For a sentence of the type "Jane did not cook the spaghetti" there are at least two possible QUDs to which the sentence answers, such us "What did Jane cook?" or "Who cooked the spaghetti?"

while for a sentence of the type “It was Jane who did not cook the spaghetti” there is only possible QUD which is “Who did not cook the spaghetti?”. So, one could hypothesize that the effects found by Tian et al. (2010, 2016) can be explainable by the ambiguity rather than the polarity of the QUD or, maybe, by a combination of the two QUD characteristics. We think that in future research should disentangle these factors: which is the role of the ambiguity of the QUD and of its polarity in the interpretation of the negation. After the results found in the two experiments here reported we think that a more direct and explicit manipulation of QUD, by inserting the negated sentence within a mini-dialog could be a good option for further studies which will try to understand the role of QUD in the online processing and in the interpretation of negative sentences.

Chapter 6: Exceptionality[6]

Introduction

In the introduction of the first experiment of this thesis we mentioned as in the literature a negative sentence is thought to be felicitous if used to describe exceptions, something that breaks the norm (Wason, 1965, Givon, 1978). Thus, the last pragmatic factor investigated in this thesis is the use of negation to describe something exceptional. In the literature it can be already found an experiment that suggests that when a negative sentence describes something exceptional its actual meaning is immediately accessed. This is a sentence verification task experiment by Arroyo (1982). In that experiment Arroyo presented participants with four words (e.g. Chicago, Miami, Paris, Dallas). Three of these four words shared a salient feature (e.g. Chicago, Miami, Dallas - they are all American cities) and one had a different feature (e.g. Paris – it is a French city). After the four words, Arroyo presented participants with a negative sentence that might be referred to one of the three words that share the feature (e.g. Dallas) or to the other word (e.g. Paris). The sentence might be true ("Dallas is not a French city"/ "Paris is not an American city") or false ("Dallas is not an American city"/ "Paris is not a French city"), and participants had to evaluate its truth value. Analyzing the reaction times in the verification task, Arroyo found that when the negative sentences were referred to one of the three words that shared the feature, participants were faster to evaluate false sentences with respect to true ones, but when they were referred to the different word, participants were faster to evaluate true sentences with respect to false ones. Of course, the result with the words that shared the feature can be interpreted as in line with the predictions of the two stage models of negation comprehension, and thus as evidence of the access to the affirmative counterpart

[6] This study has been carried out in collaboration with Matthew Crocker, Francesca Delogu, Chiara Finocchiaro, Miriam Schulz, Francesco Vespignani

meaning of the negative sentence. However, the result with the different word seems to indicate the immediate access to the actual meaning of the negation in this case. Since to our knowledge this Arroyo experiment is the only one that had showed that if a negation is used to describe something exceptional its meaning is immediately accessed, in the present experiment, we tried to find further evidence in line with this idea. To do this, we implemented a self-paced reading experiment, in which we presented negative sentences that could describe an exceptional or a non-exceptional object. Within the self-paced reading paradigm, a result that is interpreted as in line with the two-stage models of negation comprehension is that sentences that contain a negation are read slower than affirmatives, also if the two sentences are corrected for their length (Glenberg, 1999; Luedtke and Kaup, 2006, Exp.1). This result is interpreted as evidence of the additional processing step that negative sentences require with respect to affirmatives (Glenberg, 1999). At the same time, as we have seen in the second chapter of this thesis, it has been shown that it is possible to eliminate this additional processing cost for negative sentences (Glenberg, 1999; Luedtke and Kaup, 2006). This result is difficult to reconcile with the idea of a two-stages interpretation of negative sentences, since there is no evidence of the extra processing negative sentences should require with respect to affirmative ones. It follows that this result can be interpreted as evidence of the access to the actual meaning of the negation. In our experiment, in order to test if describing something exceptional enables to access immediately the actual meaning of a negative sentence, we compared the reading times of negative sentences that described something exceptional or something non-exceptional with that of their respective affirmative versions. The negative sentences started with the demonstrative pronouns “This” or “These”, in order to highlight that the sentences talked about specific items and not about prototypical items, and negated a typical property of some objects (e.g. This wedding dress is not white) or a not-typical property of the same objects (e.g. This wedding dress is not pink). This manipulation allowed us to create a condition in which the

negative sentences described exceptional objects, because they described objects without a typical property, and a condition in which they described non-exceptional objects, because they described objects without a non-typical property. The respective affirmative versions of the negative sentences were simply the same sentences without the negation (e.g. This wedding dress is white/This wedding dress is pink). We hypothesized that if describing something exceptional is actually a factor that allows to access immediately the actual meaning of negative sentences (Arroyo, 1982), participants will read negative sentences as fast as affirmative ones when negations will describe exceptional objects, but slower than affirmative ones when negations will describe non-exceptional objects. Initially this experiment was thought as an EEG experiment, but due to the lockdown situation in Germany it was modified and adapted to an on-line self-paced reading experiment.

Experiment 6

Methods

Participants

We recruited 32 participants. The sample size was exploratory. They were recruited via Prolific system (www.prolific.co) and for their participation they have been rewarded with 2.50 £. All participants declared to be between eighteen and thirty-two years old, to be German and German mother-tongue, to not have any language related disorder, to not take any drugs for mental illness currently and to not have or have had diagnosis of mental illness. All the participants that have taken part in the experiment had not taken part in two pre-tests (presented in the last section of this chapter) that we run to construct the experimental sentence stimuli of this experiment.

Stimuli

The experiment was a self-paced reading experiment. All the sentences presented were in German, half of them were affirmative and half negative. The experiment was run in German because I spent my period abroad in Saarbrucken, Germany. The sentences presented were divided in experimental and filler sentences. In experimental sentences we wanted to express the fact that we were talking about a specific object or, at least, about a specific subset of objects. To do this, all the experimental sentences presented a demonstrative pronoun (this or these) that preceded an object term (e.g. bathroom). After the object term the sentences continued describing if that object had, in case of an affirmative sentence, or did not have, in case of a negative sentence, a specific property. The property presented could be typical or not of the object mentioned. In total, in the experimental sentences, 44 different object terms were used, each with 2 different properties, one typical and one not-typical. In order to select the 44 objects terms and their properties, two pre-tests have been run, as explained in the last section of this chapter. In addition to the experimental sentences 44 filler sentences (22 affirmative and 22 negative) were constructed. Differently from experimental sentences, fillers expressed a generalization. To do this, we used objects terms in their plural form (e.g. crocodiles) in sentences that described these objects as having (or not) a property (e.g. scales). The idea to use this type of fillers, intermixed with experimental sentences, was that to bring participants to switch from sentences that express generalization to sentences that talk about a singular, specific object. This because we thought that presenting sentences that described specific objects in a context in which other sentences described a generalization could strengthen the specific interpretation of the first type of sentences. It is important to outline that both experimental and filler sentences did not terminate with the property term. Indeed, in all the sentences, after the property term the conjunction “and” (“und”) and some other words were presented before the sentences ended. For experimental sentences at least four words were

presented after the property term. Since this continuation after the property term was always affirmative, we constructed also 8 catch sentences (4 affirmative and 4 negative) which were exactly as the fillers with the only difference that after the "and" ("und") conjunction a negation (i.e. "not"/ "nicht" or "not"/"kein") was presented. With this manipulation we wanted to prevent our participants to fully predict that throughout the experiment the negation occurred only in the first part of the sentences. Lastly, for 20 experimental sentences and 24 filler sentences (12 affirmative and 12 negative) a paraphrasis of the meaning of the sentences was constructed. For half of the sentences, the paraphrasis correctly described the meaning of the sentence for the other half it did not.

Design

Our two experimental factors had two levels. Indeed, each experimental sentence could be affirmative or negative and could be about a typical or a not-typical property of one object. In this way, the experimental design was a 2(Polarity of the Sentence: Affirmative vs Negative) X 2(Typicality of the property: Typical vs Not-Typical). Since all participants went under all the experimental conditions, the design was a within-subjects. (see Table 1 for an example of experimental sentences)

	Polarity of the Sentence	**Typicality of the property**
"Dieses Hochzeitskleid ist weiß und kann mit unterschiedlichen Verzierungen geliefert warden." This wedding dress is white and can be delivered with different decorations.	Affirmative	Typical
"Dieses Hochzeitskleid ist rosa und kann mit unterschiedlichen Verzierungen geliefert warden." This wedding dress is pink and can be delivered with different decorations.	Affirmative	Not-Typical

"Dieses Hochzeitskleid ist nicht weiß und kann mit unterschiedlichen Verzierungen geliefert warden." This wedding dress is not white and can be delivered with different decorations.	Negative	Typical
"Dieses Hochzeitskleid ist nicht rosa und kann mit unterschiedlichen Verzierungen geliefert warden." This wedding dress is not pink and can be delivered with different decorations.	Negative	Not-Typical

Table 1. Example of the 4 experimental conditions for an experimental sentence.

Procedure

The stimuli of the experiment were administered via Ibex (https://spellout.net/ibexfarm, Drummond, 2013). Participants were informed that they had to read some sentences presented word by word in the centre of the screen, pressing the space bar to pass from one word to the subsequent one. They had to press the space bar with the same finger throughout the experiment. Participants were also informed that sometimes, in an unpredictable way, a paraphrasis of the sentence just read could appear, and that their task would have been to respond "True" if the paraphrasis correctly described the sentence meaning and "False" in the other case. Half of participants made True answers with their right hand and False answers with left hand, the other half made the opposite. The 176 experimental sentences constructed were divided in four different presentation lists, in a way that within one list there was no repetition of objects or properties. Each list was completed with the 44 filler and the 8 catch sentences. Each list was in turn divided in four different sub-lists of 24 sentences each. Each sub list was composed by 11 experimental sentences, 11 filler sentences and 2 catch sentences. The four sub-lists were presented in the experiment as four different experimental blocks. Half of participants saw the four experimental blocks in the ascending order (1-2-3-4) while the other half saw initially the last two blocks and then the first two (3-4-1-2). Within one

experimental block the sentences were presented in random order with the only constrain that after a maximum of three sentences without a paraphrasis, a sentence followed by a paraphrasis was presented. Each participant could see only one of the four different presentation lists. Participants were equally distributed among the lists. Before starting the experiment, in order to familiarize with the procedure, participants completed a practice block of 8 sentences.

Analysis and Results

All participants performed in the paraphrasis task with an accuracy above 80%.

We analysed reading times (RTs) of experimental sentences in four different regions of the sentence using the lme4 package (Bates, Mächler, Bolker, Walker, 2015) in Rstudio software, 1.2.1335 version (RStudio Team, 2018). The four different sentence regions analysed were: the critical word (the property), and the following three words (see Figure 1 for a general overview of the RTs results).

Reading Times

Reading Times below 50ms and above 2500ms were discarded. The resulted reading times were log-transformed with a logarithm in base 10. The log transformed reading times were analysed separately for the four regions of interest.

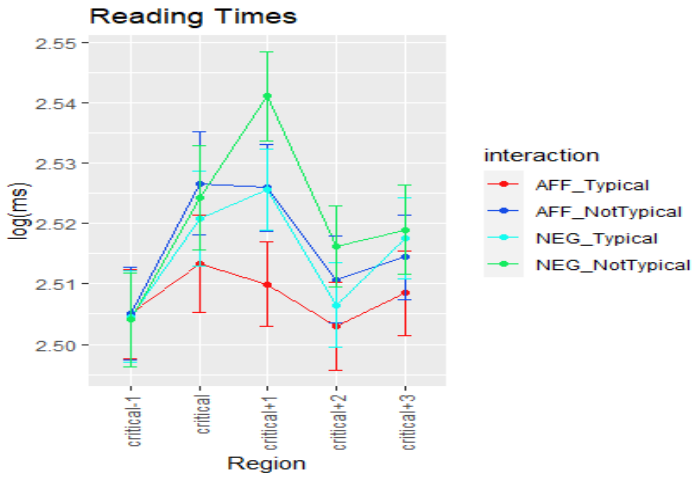


Figure 1. Mean Log-Trasformed Reading Times (corrected as in the analyses we performed).

Reading Times on the critical word

For each experimental condition, reading times that exceeded the 2.5 standard deviations above or below the mean have been removed. On the so corrected reading times five different Linear Mixed Models were fitted. The random structure of each model was composed by random intercepts for participants and items (Table 2).

Models	AIC	R Df	R Dv	Δ AIC	AICw	Df
Reading_Times ~ (1|Participants) + (1|Items)	-2503.2	1376	-2511.	2.3	0.15	4
Reading_Times ~ Polarity + (1|Participants) + (1|Items)	-2501.6	1375	-2512.	4	0.07	5
Reading_Times ~ Typicality + (1|Participants) + (1|Items)	-2505.6	1375	-2516.	0	0.47	5
Reading_Times ~ Polarity + Typicality + (1|Participants) + (1|Items)	-2503.9	1374	-2516.	1.6	0.21	6
Reading_Times ~ Polarity*Typicality + (1|Participants) + (1|Items)	-2502.6	1373	-2517.	3	0.11	7

Table 2. Models' comparison for log-transformed Reading Times. AIC = Akaike Index, R Df = Residual Degree of Freedom, R Dv = Residual Deviance, Δ AIC = delta Akaike, AICw = Akaike weight, Df = Degree of Freedom

The models' comparison among Linear Mixed Effect Models showed that the model with the main effect of Typicality was the one which better approximated the data (lower AIC index). Paired comparisons on the model selected showed that the model predicted faster reading times when the property mentioned in the sentence was the typical rather than the not-typical one (β = -0.0103, S.E. = 0.00494, z.ratio = -2.079, p = 0.0378) (see Figure 2).

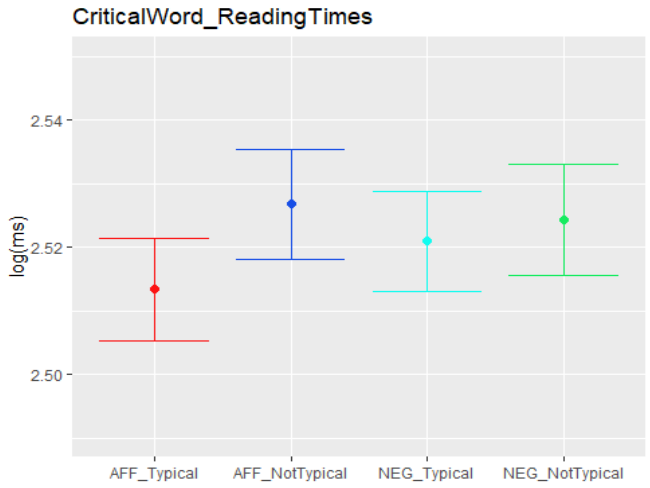


Figure 2. Mean Log-Trasformed Reading Times on the critical word.

Reading Times on the first word post-critical

For each experimental condition, reading times that exceeded the 2.5 standard deviations above or below the mean have been removed. On the so corrected reading times five different Linear Mixed Models were fitted. The random structure of each model was composed by random intercepts for participants and items (Table 3).

Models	**AIC**	**R Df**	**R Dv**	**Δ AIC**	**AICw**	**Df**
Reading_Times ~ (1|Participants) + (1|Items)	-2880.1	1370	-2888.	18.7	0	4
Reading_Times ~ Polarity + (1|Participants) + (1|Items)	-2883.7	1369	-2894.	15	0	5
Reading_Times ~ Typicality + (1|Participants) + (1|Items)	-2895.1	1369	-2905.	3.7	0.11	5
Reading_Times ~ Polarity + Typicality + (1|Participants) + (1|Items)	-2898.7	1368	-2911.	0	0.65	6
Reading_Times ~ Polarity*Typicality + (1|Participants) + (1|Items)	-2896.7	1367	-2911.	2	0.24	7

Table 3. Models' comparison for log-transformed Reading Times. AIC = Akaike Index, R Df = Residual Degree of Freedom, R Dv = Residual Deviance, Δ AIC = delta Akaike, AICw = Akaike weight, Df = Degree of Freedom

The models' comparison among Linear Mixed Effect Models showed that the additive model between Polarity and Typicality was the model which better approximated the data (lower AIC index). Paired comparisons on the model selected showed that the model predicted faster reading times when the property mentioned in the sentence was the typical rather than the not-typical one (β = -0.0176, S.E. = 0.00424, z.ratio = -4.139 , p <.0001), and when the sentence was affirmative rather than negative (β =-0.0101, S.E. = 0.00425, z.ratio = -2.379, p = 0.0175) (see Figure 3).

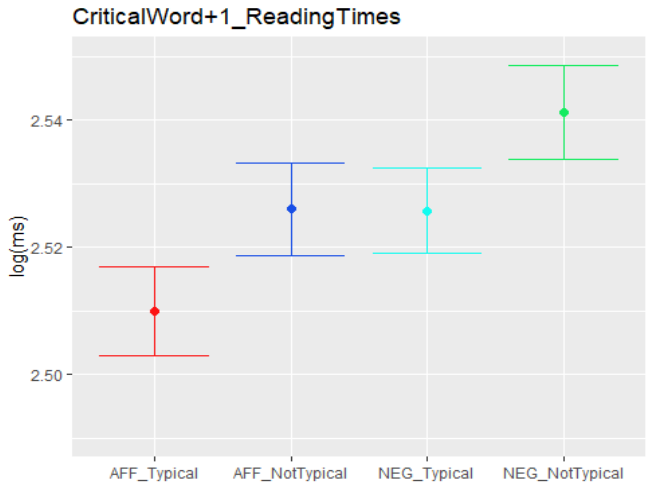


Figure 3. Mean Log-Trasformed Reading Times on the first word post critical.

Reading Times on the second word post-critical

For each experimental condition, reading times that exceeded the 2.5 standard deviations above or below the mean have been removed. On the so corrected reading times five different Linear Mixed Models were fitted. The random structure of each model was composed by random intercepts for participants and items (Table 4).

Models	AIC	R Df	R Dv	Δ AIC	AICw	Df
Reading_Times ~ (1|Participants) + (1|Items)	-2938.3	1378	-2946.	1.5	0.19	4
Reading_Times ~ Polarity + (1|Participants) + (1|Items)	-2937.0	1377	-2947.	2.8	0.1	5
Reading_Times ~ Typicality + (1|Participants) + (1|Items)	-2939.8	1377	-2950.	0	0.41	5
Reading_Times ~ Polarity + Typicality + (1|Participants) + (1|Items)	-2938.6	1376	-2951.	1.2	0.22	6
Reading_Times ~ Polarity*Typicality + (1|Participants) + (1|Items)	-2936.6	1375	-2951.	3.2	0.08	7

Table 4. Models' comparison for log-transformed Reading Times. AIC = Akaike Index, R Df = Residual Degree of Freedom, R Dv = Residual Deviance, Δ AIC = delta Akaike, AICw = Akaike weight, Df = Degree of Freedom

The models' comparison among Linear Mixed Effect Models showed that the model with the main effect of Typicality was the one which better approximated the data (lower AIC index). Paired comparisons on the model selected showed that the model predicted a non-significant

trend towards faster reading times when the property mentioned in the sentence was the typical rather than the not-typical one (β = -0.00794, S.E. = 0.00423, z.ratio = -1.876, p = 0.0609) (see Figure 4).

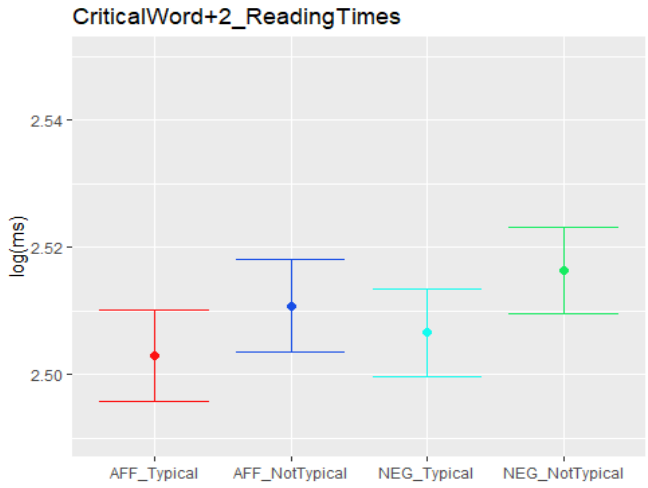


Figure 4. Mean Log-Trasformed Reading Times on the second word post critical.

Reading Times on the third word post-critical

For each experimental condition, reading times that exceeded the 2.5 standard deviations above or below the mean have been removed. On the so corrected reading times five different Linear Mixed Models were fitted. The random structure of each model was composed by random intercepts for participants and items (Table 5).

Models	AIC	R Df	R Dv	Δ AIC	AICw	Df
Reading_Times ~ (1|Participants) + (1|Items)	-2987.7	1380	-2996.	0	0.32	4
Reading_Times ~ Polarity + (1|Participants) + (1|Items)	-2986.7	1379	-2997.	0.9	0.2	5
Reading_Times ~ Typicality + (1|Participants) + (1|Items)	-2987.3	1379	-2997.	0.4	0.26	5
Reading_Times ~ Polarity + Typicality + (1|Participants) + (1|Items)	-2986.3	1378	-2998.	1.3	0.16	6
Reading_Times ~ Polarity*Typicality + (1|Participants) + (1|Items)	-2984.3	1377	-2998.	3.3	0.06	7

Table 5. Models' comparison for log-transformed Reading Times. AIC = Akaike Index, R Df = Residual Degree of Freedom, R Dv = Residual Deviance, Δ AIC = delta Akaike, AICw = Akaike weight, Df = Degree of Freedom

The models' comparison among Linear Mixed Effect Models showed that the null model was the one which better approximated the data (lower AIC index) (see Figure 5).

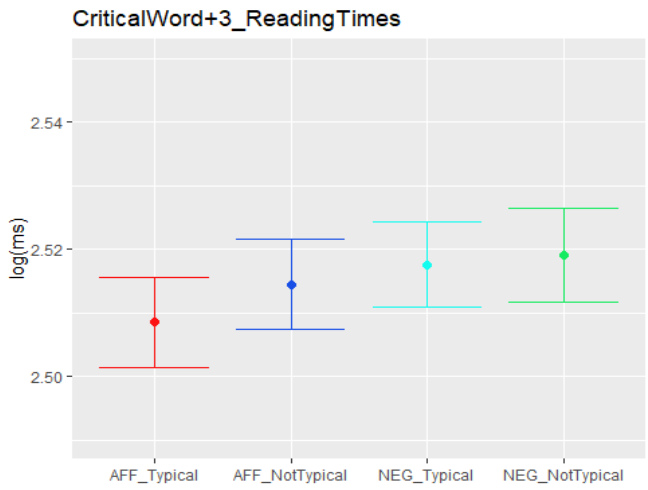


Figure 5. Mean Log-Trasformed Reading Times on the third word post critical.

Discussion

The aim of the present experiment was to find evidence in line with the idea that describing an exception is one of those pragmatic factors that allows to access immediately the actual meaning of a negative sentence. To do this, we implemented a self-paced reading experiment in which we presented negative sentences that described exceptional objects and negative sentences that described non-exceptional objects. In details, negative sentences that described exceptional objects were about some objects without one their typical property, while negative sentences that described non-exceptional objects were about some objects without one their not-typical property. Reading times of these negative sentences have been compared with reading times of the respective affirmative versions. Our main hypothesis was that, if describing something exceptional allows to access immediately the actual meaning of a negative sentence (Arroyo, 1982), participants would have read negative sentences that described an exceptional object as fast as their respective affirmative versions, but slower than

their respective affirmative versions when they described a non-exceptional object. A result of this type could be taken as evidence that if a negative sentence describes something exceptional does not require any extra processing with respect to its affirmative counterpart. Of course, this kind of interpretation would be at odds with the view of the two-stage models of negation comprehension. Our analysis, however, revealed a main effect of the negation that indicated that both if the negative sentences described an exceptional or a not-exceptional object, they were read slower with respect to affirmative ones. Of course, this result is against our hypothesis and so does not allow us to conclude that if negation is used to describe something exceptional its actual meaning is immediately access. At the opposite, a result of this type, can be interpreted as in line with the predictions of the two-stage models of negation comprehension, that always assume an extra processing for negative sentences with respect to affirmative ones. In this way, our result seems to challenge the idea of an immediate access to the actual meaning of the negative sentence in the case in which negation described something exceptional and, for this reason, seem to be in contrast with the results by Arroyo (1982). We think, anyway, that there is a major difference between our experiment and that of Arroyo (1982), that can explain the contrasting results. Indeed, while in the experiment by Arroyo participants could predict what would have been the negated word, in our experiment they could not. Indeed, if in our experiment for a sentence as "This wedding dress is not…" there can be different plausible endings (e.g. "tight-fitting", "white", "expansive", "low-cut", "pink" etc.), in the experimental environment of Arroyo for a sentence as "Paris is not …" that followed a context of the type "Chicago, Miami, Paris, Dallas", "American city" was highly the most plausible continuation. We think that this difference in the predictability of the negated word can account for the contrasting results between our and Arroyo experiment. It is at the same time obvious that prediction cannot explain entirely the results by Arroyo (1982), otherwise he would have found evidence for an access to the actual meaning of the negation

also in the case in which negative sentences were used to refer about the non-exceptional word (e.g. Dallas). We think, indeed, that it is the combination of both pragmatic factors (exceptionality) and predictability that can explain the results by Arroyo (1982), not only one of the two factors (this idea will be further discuss in the Chapter 7).

Construction of the Experimental Sentences

In this section we report the procedure followed to construct the experimental sentences of the self-paced reading experiment.

First Pre-Test: Typicality Test

We run a first pre-test with the goal to select a series of objects that had at least one property that was judged as typical and one that was judged as not-typical of the object. To do this, to 79 different objects terms (e.g. bathroom), we associated at least 2 properties, one which we thought was typical (e.g. WC) and one which we thought was not-typical (e.g. heated floor) of the object. A maximum of 6 different properties for a single object have been proposed. For each object-property pair generated, a question that asked how typical was for that object to have that specific property was developed (e.g. "Wie üblich ist es für ein Badezimmer, eine Bodenheizung zu haben?"/ "How typical is for one bathroom to have a heated floor?"). In total 344 questions were developed. The 344 questions have been split in six different presentation lists, in a way that within one list there were no repetitions of objects and properties. Since some presentation lists had less questions than the other lists and since within each list the number of questions about typical and not-typical properties was unbalanced, we developed 19 filler questions. These filler questions were added to each presentation list until the number of typical and not-typical properties presented within each list was equal and until each list was composed by 76 questions. Of course, also after adding the fillers within each list there was

not any repetition for objects or properties. For each presentation list we recruited 10 participants, thus 60 participants were recruited in total. Participants' task was simply to read the questions and to provide an answer by using a seven - points scale that went from 1) not typical (nicht üblich) to 7) very typical (sehr üblich). Participants were recruited via Prolific system (www.prolific.co) and for their participation they have been rewarded with 1.25£. All participants declared to be between eighteen and thirty-two years old, to be German and German mother-tongue, to not have any language related disorders. Within each presentation list, the questions were presented in random order. Their presentation was controlled through Ibex farm (https://spellout.net/ibexfarm, Drummond, 2013). Before starting the pre-test participants familiarized with the task responding to six practice questions. Finally, in order to select the objects that met our requirements, we decided to consider as typical properties of the object only the ones that received a mean-rating equal or above 5.5, and as not-typical the ones that received a mean-rating equal or below 4. Moreover, for the same object, we decided that between the typical and the not-typical property selected there should be a difference of 2 points at least. In this way, of the 79 original objects, 68 different objects, that had at least one typical and one not typical property, were selected.

Second Pre-Test: Existence Test

The 68 objects selected in the first pre-test experiment went under a second pre-test. The aim of this second pre-test was to select only the objects that with their not-typical property (e.g. Banana – brown) and without their typical property (e.g. Banana - not yellow) were judged as existing. We decided to run this second pre-test because we wanted to avoid in our self-paced reading experiment to present sentences about objects considered as non-existent. To do this, for each of the 68 objects, we selected a maximum of four properties from that selected in the previous pre-test, of which at least one typical and one not-typical property. For each of the selected object-property pairs, we constructed a question. If the property was not-typical the

question asked if that object exists in the real world with that specific property (e.g. “Is there a situation in the real world in which one banana is brown?”/ “Gibt es in der realen Welt eine Situation, in der eine Banane braun ist?”). If the property was typical, the question asked if that object exists in the real world without that specific property (“Is there a situation in the real world in which one banana is not yellow?”/ “Gibt es in der realen Welt eine Situation, in der eine Banane nicht gelb ist?”). Since in the questions asking whether the objects existed with their not-typical property object and property terms were always weakly related whereas in the questions asking whether the object existed without their typical property object and property terms were always highly related, we decided to construct also questions asking whether the same object existed with its typical property (“Is there a situation in the real world in which one banana is yellow?”/ “Gibt es in der realen Welt eine Situation, in der eine Banane gelb ist?”) or without its not-typical property (e.g. “Is there a situation in the real world in which one banana is not brown?”/ “Gibt es in der realen Welt eine Situation, in der eine Banane nicht braun ist?”). All the questions generated were divided in 8 different presentation lists, in a way that within one list there were no repetitions for objects or properties. In the resulted lists there were 54 questions per lists, and there was a little unbalance with thirty questions (~55%) about not-typical properties and twenty-five questions (~45%) about typical properties. Since the difference was so little, we decided to not correct it. Moreover, since potentially all our questions could report a positive answer, we constructed also 54 filler questions (27 that asked if one object exists with a property, and 27 that asked if one object exists without a property) that required a negative answers (e.g. “Is there a situation in the real world in which a sheep can talk?”/ “Gibt es in der realen Welt eine Situation, in der ein Schaf das sprechen kann?”). For each presentation list we recruited 10 participants, so 80 participants were recruited in total. Participants that had taken part in this pre-test did not take part in the previous one. Participants were recruited on Prolific (www.prolific.co) and for their participation have been rewarded with

1.25£. All participants declared to be between eighteen and thirty-two years old, to be German and German mother-tongue, to not have any language related disorders. Questions within each list were presented in random order and their presentation was controlled via Ibex farm (https://spellout.net/ibexfarm, Drummond, 2013). Participants' task was to read the questions and to provide an answer by using a five points scale (1) no, 2) I don't think so, 3) I don't know, 4) I think so, 5) yes). To familiarize with the task, participants were required to answer to eight questions of practice before starting the pre-test. Moreover, to ensure participants were paying attention while answering to the questions, four attention checks were presented during the pre-test. We decided to evaluate as existing objects only objects that with their not-typical and without their typical property received a "yes" or a "think yes" answer from at least seven participants out of ten. In this way sixty objects, out of the sixty-eight original ones, were selected.

Final selection for experimental sentences

From the 60 objects selected in the second pre-test we selected 44 objects, with one typical and one not-typical property. These 44 objects were selected to be used in the experimental sentences of the self-paced reading experiment. The 16 objects we did not select were discarded because we were not able to create with them a continuation of the sentences that after the "and" ("und") conjunction was the same for all the experimental conditions (Polarity of the sentence: affirmative vs negative, Typicality: typical vs not-typical) in which that object was mentioned. Indeed, in the experiment we wanted that for each object, after the "and" ("und") conjunction the sentences terminated with the same words in all the experimental conditions (see Table 1).

Chapter 7: General Discussion and Conclusion

General Discussion

In the first chapter(s) of this thesis, we have reviewed a number of empirical results in line with the proposal of the two-stage models of negation interpretation. These models propose that the actual meaning of a negative sentence, differently from that of an affirmative sentence, cannot be accessed immediately (incrementally) but only after its affirmative counterpart meaning is computed. In other words, these models propose that to understand "the boy is not eating" the system must pass through a first moment in which the meaning "the boy is eating" is accessed. Interestingly, there is also evidence suggesting that in some situations it is possible to directly access the actual meaning of a negative sentence, incrementally and with no delay. It has been proposed that this may happen when negative sentences are uttered and interpreted within pragmatic contexts that make them felicitous. Thus, pragmatic factors would be responsible for the immediate access to the actual meaning of a negative sentence. Despite the undoubtful interest of this proposal, it is to date not entirely clear what makes a negative sentence felicitous. The aim of this thesis was to shed light on the pragmatic factors that may allow an immediate access to the actual meaning of the negation. To do this we run six different experiments, using different experimental paradigms, in which we investigated four different pragmatic factors: 1) the fact that the negation is used to refer to a minority, 2) the fact that the negation describes something that non-exists, 3) the fact that the negation answers to a negative Question Under Discussion (QUD), 4) the fact that the negation is used to refer about an exception. For the minority factor, we implemented a sentence-picture verification task in which we presented affirmative and negative sentences that could refer to a majority or to a minority of objects depicted in the picture. In sentence-picture verification tasks, it is often found an interaction

between polarity and truth-value showing that with affirmative sentences participants are facilitated in evaluating "true" sentences with respect to "false" ones whereas with negative sentences they are facilitated in evaluating "false" sentences with respect to "true" ones. This result is taken as evidence in line with the proposal of the two-stage models of negation interpretation. In our experiment we investigated whether the reference to a minority leads to facilitation on negative true sentences with respect to false ones, thus eliminating the polarity by truth-value interaction effect. Of course, a result of this type could have been interpreted as against the proposal of the two-stage models and in line with the idea of an immediate access to the actual meaning of the negative sentences referred to a minority. Contrary to our predictions, however, both when the sentences referred to the majority or to the minority, we found the polarity by truth-value interaction effect. This result shows that the reference about a minority is not, at least in our paradigm, a sufficient pragmatic factor to allow immediate access to the actual meaning of the negative sentence.

To test the nonexistence factor, we run other two sentence-picture verification experiments. We compared one negative sentence used to describe something that non-exists with one negative sentence used to describe an alternative. The difference between the two experiments was that, in the first one, participants could fully predict the negated word presented in the sentences while, in the second one, two possible words could complete the sentences with the same probability. Our prediction was to eliminate the polarity by truth-value interaction effect in the case in which the negative sentence described something that non-exists, and so to find that in this case participants were facilitated in evaluating true sentences vs false ones. The results of both experiments indicated that with the negation that described something that non-exists we actually found participants were facilitated in evaluating true sentences vs false ones while in the case of the negation that described an alternative we found a speed-accuracy trade off, with faster but less accurate answers with true with respect to false sentences. The

comparison of the nonexistence negation with the alternative one ensured that, in the case we would have found the elimination of the polarity by truth-value interaction effect for the non-existence negation, the elimination could be entirely attributable to the fact that the negation was used to describe something that non-exists and not to other factors. Unfortunately, the speed-accuracy trade-off for the alternative negation does not allow us to say if also in the alternative condition participants were able to immediately access the actual meaning of the negation. Consequently, there is no conclusive evidence that describing something that non-exists is a key pragmatic factor enabling the immediate access to the actual meaning of the negation.

The third factor we tested has been proposed by Tian and colleagues (2010) and is the fact that a negative sentence answers to a negative Question Under Discussion (QUD). As we have seen in Chapter number 5, in a theoretically relevant experiment, Tian et al. (2010) found evidence in line with the idea that if a negative sentence answers a negative QUD its actual meaning is immediately accessible but if it answers a positive QUD its affirmative counterpart meaning is initially accessed. In a first experiment we tried to replicate this result within a sentence-picture verification task in which we compared negative sentences that answered a negative QUD with negative sentences that answered a positive QUD. Of course, we predicted to eliminate the polarity by truth-value interaction effect in the case in which the negative sentences answered a negative QUD. However, contrary to our predictions, we found the classical polarity by truth-value interaction effect for negative sentences that answered either a negative QUD or a positive QUD. The result is of course at odds with the proposal and the results by Tian at al. (2010) because it seems to indicate that the fact that a negative sentence answers to a negative QUD is not a factor that allows to access immediately the actual meaning of the negation. However, since there were many differences between the paradigm we used in our experiment and the paradigm Tian and colleagues used in their original experiment (a probe-recognition

task), we decided to run a faithful replica in Italian of their experiment in order to understand if our replication failure was due to the experimental method we used. In probe-recognition tasks it is often found that with negative sentences participants are faster to respond to pictures that mismatch the sentence meaning with respect to pictures that match it, a result that is interpreted as in line with the two-stage models of negation interpretation. In their experiment Tian et al. (2010) were able to reverse this effect for the negative sentences that answered a negative QUD but they found it for negative sentences that answered a positive QUD. In our faithful replica, however, we have not been able to replicate their results since we did not find any difference in the response times between pictures that matched or mismatched the sentence meaning. This null result is at odd with the proposal by Tian et al. (2010), but given that we did not find the classical effect of faster responses to picture that mismatched the sentence meaning with respect to picture that matched it, neither for negative sentences that answered a positive QUD nor for negative sentences that answered a negative QUD, the null result would mainly speak against robustness and reliability of the experimental paradigm rather than against the theoretical proposal of a key role of polarity of QUD in making negative sentences felicitous. In order to account for the discrepancy between our results and Tian et al. (2010) findings, we proposed that our failure to get any effect could be attributed to the online modality of our experiment (among other possible other causes discussed in detail in chapter 5).

The fourth factor we tested was the use of the negation to describe something exceptional. In a truth-judgment task experiment, Arroyo (1982) found evidence in line with the idea that describing something exceptional is a factor that allows to access immediately the actual meaning of the negation. However, since to our knowledge the result has never been replicated, we thought it would be important to verify this finding. So, in our last experiment, we implemented a self-paced reading experiment in which we compared the reading times of negative sentences with that of their respective affirmative versions. The negative sentences

could describe something exceptional (e.g. "*This wedding dress is not white"*) or something non-exceptional (e.g. "*This wedding dress is not pink*"). Usually, in self-paced reading experiments a result that is taken as evidence in support of the two-stage models of negation interpretation is the fact that negative sentences are read slower with respect to their affirmative versions. Some studies, however, have found that this cost can be eliminated. Of course, the results of these studies are not in line with the proposal of the two-stage models and can be taken as evidence of an immediate access to the actual meaning of negative sentences. Thus, we hypothesized that if describing something exceptional allows to immediately access the actual meaning of the negation, participants would have read negative sentences as fast as affirmative ones in the exceptional condition. Our results, however, indicated that independently from the condition (exceptional vs. non-exceptional) participants read the negative sentences slower than affirmative sentences. As a consequence, we cannot conclude that describing something exceptional is a pragmatic factor that allows to immediately access the actual meaning of the negation. As it is clear from this summary, for all the four pragmatic factors we tested we have never found clear evidence that specific key pragmatic factors allow the immediate access to the actual meaning of the negation. However, we think that different considerations can be drawn for the different factors we considered. Indeed, for the minority factor our results seem to be quite clear in indicating that, at least within our paradigm, this is not a factor that allows to access immediately the actual meaning of the negation. Instead, for the nonexistence factor the results are not so clear. Indeed, in our experiments we found evidence in line with the idea that when a negative sentence described something that non-exists its actual meaning was immediately accessed but, at the same time, we found a speed-accuracy trade off when the negative sentence described an alternative, leaving open the hypothesis that immediate access may have been there also in this case. Since we cannot exclude this thing, we cannot rule out the possibility that methodological factors of our

experiments, such as for example the possibility to predict the negated word presented in the sentences with at least 50% of accuracy and/or the possibility to predict the affirmative meaning of the negative sentences (e.g. “not black” always meant “white” in this experiment), are factors that allow to access immediately the actual meaning of the negation. As a consequence, we cannot keep out the possibility that the access to the actual meaning we found with the nonexistence negation was due to these methodological/predictability factors rather than to the fact that the negation was used to describe something that non-exists. At the same time, however, one can speculate that if these methodological/predictability factors were actually factors that allow to access immediately the actual meaning of the negation we should have found clear evidence of this access (and not a speed-accuracy trade off) also in the case in which the negative sentence described an alternative. So, following this speculation, we propose that probably the fact that the negation is used to describe something that non-exists affects the likelihood to access immediately the actual meaning of the negation. In light of this, we think it would be worth to further study the nonexistence factor trying to disentangle its effect from the predictability effects that may have a key role in eliminating the polarity by truth interaction in our experiments. As to the answer to a negative QUD factor, our sentence-picture verification task seems to speak against the proposal of Tian et al. (2010), and so against the idea that this is a factor that enable to access immediately the meaning of a negative sentence. On the other hand, the results of the probe-recognition task are not incompatible with this idea, though they not directly support it. As written above, we think that the result of the probe-recognition task, being a null result, cannot be taken as a falsification of the proposal but again, as an indication of the weakness of the paradigm. For the sentence-picture verification task we think that methodological aspects that differentiate our experiment from the experiments by Tian et al. (2010, 2016) can potentially explain why we failed to find evidence in support of their proposal. First of all, we have proposed that possibly the massive repetition

of same lexical items referring to a rather abstract and impoverished semantic world in our experiment vs the use of more naturalistic situations which varied from trial to trial in their experiments (2010, 2016) could have discouraged our participants to derive a negative QUD in the environment of our experiment. Secondly, we think that two other valid reasons that can explain our failure can be the explicit truth-judgment task that was required in our experiment, but that was not required in the experiments by Tian and colleagues (2010, 2016), and the different negated word presented in our experiment vs Tian et al. experiments (2010, 2016). Indeed, in both their experiments, Tian et al. (2010, 2016) presented binary predicates (e.g. open/closed, switched off/switched on) while in our experiment we presented unary predicates (e.g. red, green, etc.). Binary predicates have the advantage that when they are negated their affirmative meaning is predictable, "not closed" means "open", while for unary predicates this is not possible, "not red" might mean yellow, blue, green, etc.. With our experiment we are not able to disentangle which is the reason for which we did not find evidence il line with the Tian's proposal, but our results clearly show that the effect of the polarity of QUD manipulated using cleft sentences does not suffice by itself to find evidence of the immediate access to the actual meaning of negations. Anyway, as we have written in the discussion of the Chapter 5 we think that our results cannot be taken as a falsification of the Tian et al. (2010) proposal that, at the opposite, we think it is worth to be further studied, possibly manipulating more explicitly QUD within short discourse contexts and investigating other dimensions of the QUD rather than only its polarity. Even for the last factor we tested, the exceptionality factor, we did not find evidence in line with the idea of an immediate interpretation of the negation. Also in this case, as for the QUD, in the literature can be found one experiment that seems to indicate that this is a factor that allows to access immediately the actual meaning of the negation (Arroyo, 1982). There are some major differences between our experiment and that of Arroyo that can explain our failure in finding evidence of the immediate access to the actual meaning of the negation. First of all,

we used a self-paced reading paradigm while Arroyo used a truth-judgment paradigm, so it can be hypothesized that the effect of exceptionality is difficult to be generalized to self-paced reading task. In addition, the predictability of the negated word presented in the sentences was much higher in Arroyo's experiment than in our experiment (see the discussion of the Chapter 6). It is interesting to note that the predictability factor seems to be a constant difference between our experiments and the ones that in the literature, studying the same pragmatic factors, found evidence for the access to the actual meaning of the negation. Probably, the relationship between predictability and the immediate access to the actual meaning of the negation is more complex than believed and needs further investigation for its potential interference with other pragmatic factors (see also chapter II). Indeed, looking at the experiments that found evidence of the immediate access to the actual meaning of the negation, one can see how all these experiments presented at least one of these predictability effects. Indeed, in Nieuwland and Kuperberg (2008), Dale & Duran (2011), Tian et al. (2010, 2016) experiments, the authors presented sentences with binary predicates that in this way enable to predict the affirmative meaning of the negation. On the other hand, the study by Arroyo (1982) allowed to predict with high accuracy the word that would have been presented. Looking at these experiments one could conclude that predictability factors seem to be necessary in order to find evidence in line with the idea of an immediate access to the actual meaning of the negation. A similar conclusion seems can be drawn also from our experiments. Indeed, the only condition in in which we found evidence for an immediate access to the actual meaning of the negative sentences is the condition in which we presented a negative sentence to describe something that non-exists, in the two experiments in which we tested the nonexistence factor. As written above, in these experiments our participants could predict (fully or with a 0.5 uncertainty) the negated word and the affirmative meaning of the negative sentences. At a theoretical and linguistic level predictability is clearly an aspect which is not independent from

pragmatic felicity. For example, within Grice approach to define a speech act felicitous the contribute to the conversation should be in line with what expected by the interlocutor, the theoretical framework of QUD is linked to specific expectation of the addressee about what will follow in the message. The cloze probability procedure, typically used in psycholinguistics to operationalize predictability of a word in a sentence context was developed by Taylor (1953) with the aim of measuring text readability, within the assumption that in a (pragmatically) coherent, easy to read, text any single word could be predicted on the basis of the discourse context. However, despite the growing role of predictions and expectations in the modelling of sentence comprehension (Pickering and Garrod, 2013) and, more in general, cognition (Bar, 2009) and neuroscience (Friston, 2010), we think that predictability and expectations are still rather vaguely conceptualized in terms of clear psychological constructs. Moreover, in our specific field it would be important to link psychological constructs linked to prediction with linguistic concepts linked to prediction (such as the QUD framework in pragmatics). In order to do so it would be important to detail accurately what predictability means. Of course, this is not an easy task at all. Indeed when predictability refers to words, it may mean many things, for example: we can predict the specific word at abstract or at orthographic/phonological level, we can predict some properties of an expected word (such as word category or specific features like gender and number), we can predict the broader meaning of the word, we can predict a general situation that makes some words more expected than others, etc.. For sure it is difficult to define these relations, however, we think that paradigms similar to the sentence-pictures verification tasks we implemented in this thesis could help in the control or manipulate many of these predictability variables. Indeed, after defining which type of predictability we are talking about, one can have full control on this variable. For example, in our experiments we had full control on how much the specific negated word we presented was predictable (100% when we presented only one negated word, 50% when we presented two negated words, etc.),

and also how much the affirmative meaning of the negative sentence was predictable (100% when there was only one affirmative meaning for a negative sentence, 50% if the negative sentence can have two different meaning, etc.). In this way, the adoption of these types of paradigms seems that could help in understanding the specific weight that predictability and pragmatic factors could have in the possibility to access immediately the actual meaning of a negative sentence. Of course, despite the advantage that an approach of this type seems to have within experimental settings, an open question remains if results from these paradigms can be generalized to more natural and realistic language use situations.

Conclusion

To conclude, after three years of uninterrupted studies of negation comprehension, the main conclusion is apparently unsatisfactory in that I think still much work needs to be done before having a full comprehension of the phenomenon. In this thesis we focused in trying to identify which are pragmatic factors that allow to access immediately the actual meaning of the negation since an abundant literature seem to indicate that, out of a pragmatic context that make them felicitous, negative sentences are understood in two stages. As it is clear from the General Discussion, however, we failed in identifying just one of these pragmatic factors, and in addition our results seem to suggest that only in really specific situations can be found evidence of the immediate access to the actual meaning of the negation, in particular, when these "formal" pragmatic properties are played in a paradigm in which predictability matters. In this way our results seem to support the idea of the two-stages models of negation interpretation and so that generally negation is understood in two steps, at least in the laboratory environment in which different subsequent trials are to be interpreted as out of the blue. Clearly, in realistic language use situations, such conversations or connected long text context, discourse representations support stronger expectations and lager predictability than single sentence in insolation, presented in a lab. Anyway, I think that is worth to mention that, despite the fact that the idea that out of an adequate pragmatic context negative sentences are interpreted in two stages seem to have many evidence in support, in the literature can be found a body of recent evidence that seem to suggest that negation can be immediately taken into consideration during the interpretation of the sentence, also when negative sentences are presented out of an adequate pragmatic context. This evidence came mainly from neuroscience studies (Tettamanti, Manenti, Della Rosa, Falini, Perani, Cappa & Moro, 2008, Tomasino, Weiss and Fink, 2010 and Papeo, Hochmann, & Battelli 2016). It is clear that an idea of this type is completely against the view of the two-stages models and seems to be also in odds with the

results we found in our experiments. However, it is important to outline the radical shift of perspective that a view of this type, we think, would have on the research of negation comprehension. Within this view, indeed, it would not make any sense to search pragmatic factors that allow to access immediately the actual meaning of the sentence but, at the opposite, it would be really important to define accurately which are, or which is, the condition(s) that bring a two stages interpretation of the negation (condition(s) that should be met also in our experiments). Indeed, we firmly think that only when we would have rigorously defined when negation is understood in one or two-stages, solid theories about how negation is understood can be developed. It is important to outline that we are not theoretically biased, and so we do not have theoretical reasons to believe that one perspective that think negation is usually interpreted in two-stages is intrinsically better or worse than a perspective that think negation is usually interpreted in one-stage. Anyway, we think that it is worth mentioning a potential conflict between the two-stages accounts of negation comprehension and the general accepted idea of language comprehension as incremental (Altmann & Mirković, 2009; Pickering & Garrod, 2007). Incrementality of language comprehension means that every word is comprehended exactly in the moment at which it is read or heard. Of course, a view of negation interpretation that assumes that after reading a negative sentence the negation is not immediately considered because its affirmative counterpart meaning is initially accessed is at odds with the idea that language comprehension is incremental. So, the only way to accept the two-stage models within this framework would be to consider negation as an exceptional case in language comprehension. This is a licit move but, following the reasoning of Occam's razor, it would be easier to consider negation as all the other words. Anyway, as written above, we think that new data and much research is needed to arrive to the conclusion if negation is like all the other words or not.

References:

Anderson, S. E., Huette, S., Matlock, T., & Spivey, M. J. (2010). On the temporal dynamics of negated perceptual simulations. *Meaning, form, & body*, 1-20.

Arroyo, F. V. (1982). Negatives in context. *Journal of Verbal Learning and Verbal Behavior*, *21*(1), 118-126.

Altmann, G. T., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition*, *73*(3), 247-264.

Altmann, G. T., & Mirković, J. (2009). Incrementality and prediction in human sentence processing. *Cognitive science*, *33*(4), 583-609.

Altmann, G. T., & Kamide, Y. (2007). The real-time mediation of visual attention by language and world knowledge: Linking anticipatory (and other) eye movements to linguistic processing. *Journal of memory and language*, *57*(4), 502-518.

Bahlmann, J., Mueller, J. L., Makuuchi, M., & Friederici, A. D. (2011). Perisylvian functional connectivity during processing of sentential negation. *Frontiers in Psychology*, *2*, 104.

Bar, M. (2009). Predictions: a universal principle in the operation of the human brain.

Baroni, M., Bernardini, S., Ferraresi, A., & Zanchetta, E. (2009). The WaCky wide web: a collection of very large linguistically processed web-crawled corpora. *Language resources and evaluation*, *43*(3), 209-226.

Baroni, M., & Kilgarriff, A. (2006). Large linguistically-processed web corpora for multiple languages. In *Demonstrations*.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. *arXiv preprint arXiv:1406.5823*.

Blank, M. A. (1980). Measuring lexical access during sentence processing. *Perception & Psychophysics*, *28*(1), 1-8.

Blasko, D. G., & Connine, C. M. (1993). Effects of familiarity and aptness on metaphor processing. *Journal of experimental psychology: Learning, memory, and cognition*, *19*(2), 295.

Calignano, G. (2020). Linguistic information and visual attention deployment: the in uence of meaningful labels on the orienting of attention. PhD dissertation, Trento University, Rovereto (TN), Italy

Carpenter, P. A., & Just, M. A. (1975). Sentence comprehension: a psycholinguistic processing model of verification. *Psychological review*, *82*(1), 45.

Carpenter, P. A., Just, M. A., Keller, T. A., Eddy, W., & Thulborn, K. (1999). Graded functional activation in the visuospatial system with the amount of task demand. *Journal of Cognitive Neuroscience*, *11*(1), 9-24.

Christensen, K. R. (2009). Negative and affirmative sentences increase activation in different areas in the brain. *Journal of Neurolinguistics*, *22*(1), 1-17.

Clark, H. H., & Chase, W. G. (1972). On the process of comparing sentences against pictures. *Cognitive psychology*, *3*(3), 472-517.

Clark, H. H., & Chase, W. G. (1974). Perceptual coding strategies in the formation and verification of descriptions. *Memory & Cognition*, *2*(1), 101-111.

Colston, H. L. (1999). "Not good" is "bad," but "not bad" is not "good": An analysis of three accounts of negation asymmetry. Discourse Processes, 28(3), 237-256.

Cooper, R. M. (1974). The control of eye fixation by the meaning of spoken language: a new methodology for the real-time investigation of speech perception, memory, and language processing. *Cognitive psychology*.

Cornish, E. R., & Wason, P. C. (1970). The recall of affirmative and negative sentences in an incidental learning task. *Quarterly Journal of Experimental Psychology*, *22*(2), 109-114.

Dale, R., & Duran, N. D. (2011). The cognitive dynamics of negated sentence verification. *Cognitive science*, *35*(5), 983-996.

Darley, E. J., Kent, C., & Kazanina, N. (2020). A 'no'with a trace of 'yes': A mouse-tracking study of negative sentence processing. *Cognition*, *198*, 104084.

de Koning, B. B., Wassenburg, S. I., Bos, L. T., & Van der Schoot, M. (2017). Size does matter: Implied object size is mentally simulated during language comprehension. *Discourse processes*, *54*(7), 493-503.

Drummond, A. (2013). Ibex farm. *Online server: http://spellout. net/ibexfarm*, *831*.

Dudschig, C., Mackenzie, I. G., Maienborn, C., Kaup, B., & Leuthold, H. (2019). Negation and the N400: investigating temporal aspects of negation integration using semantic and world-knowledge violations. *Language, Cognition and Neuroscience*, *34*(3), 309-319.

Fischler, I. (1977a). Associative facilitation without expectancy in a lexical decision task. *Journal of Experimental Psychology: Human Perception and Performance*, *3*(1), 18.

Fischler, I. (1977b). Semantic facilitation without association in a lexical decision task. *Memory & cognition*, *5*(3), 335-339.

Fischler, I., Bloom, P. A., Childers, D. G., Roucos, S. E., & Perry Jr, N. W. (1983). Brain potentials related to stages of sentence verification. *Psychophysiology*, *20*(4), 400-409.

Fischer, M. H., & Hartmann, M. (2014). Pushing forward in embodied cognition: may we mouse the mathematical mind?. *Frontiers in psychology*, *5*, 1315.

Freeman, J. B., & Ambady, N. (2010). MouseTracker: Software for studying real-time mental processing using a computer mouse-tracking method. *Behavior research methods*, *42*(1), 226-241.

Friston, K. (2010). The free-energy principle: a unified brain theory?. *Nature reviews neuroscience*, *11*(2), 127-138.

Ganis, G., Kutas, M., & Sereno, M. I. An Electrophysiological Study of the Comprehension of Words and Pictures in Reading. *Journal of Cognitive Neuroscience*, *82*, 89-106.

Giora, R., Balaban, N., Fein, O., & Alkabets, I. (2004). Explicit negation as positivity in disguise. In *Figurative language comprehension* (pp. 245-270). Routledge.

Givón, T. (1978). Negation in language: pragmatics, function, ontology. In Pragmatics (pp. 69-112). Brill.

Glenberg, A. M., Robertson, D. A., Jansen, J. L., & Johnson-Glenberg, M. C. (1999). Not propositions. *Cognitive Systems Research*, *1*(1), 19-33.

Glucksberg, Sam, Newsome, Mary R., Goldvarg, Yevgeniya, 2001. Inhibition of the literal: filtering metaphor-irrelevant information during metaphor comprehension. Metaphor and Symbol 16, 277–294

Gough, P. B. (1965). Grammatical transformations and speed of understanding. *Journal of verbal learning and verbal behavior*, *4*(2), 107-111.

Gough, P. B. (1966). The verification of sentences: The effects of delay of evidence and sentence length. *Journal of verbal learning and verbal behavior*, *5*(5), 492-496.

Hasegawa, M., Carpenter, P. A., & Just, M. A. (2002). An fMRI study of bilingual sentence comprehension and workload. *Neuroimage*, *15*(3), 647-660.

Hasson, U., & Glucksberg, S. (2006). Does understanding negation entail affirmation?: An examination of negated metaphors. *Journal of Pragmatics*, *38*(7), 1015-1032.

Herbert, C., & Kübler, A. (2011). Dogs cannot bark: Event-related brain responses to true and false negated statements as indicators of higher-order conscious processing. *PloS one*, *6*(10), e25574.

Horn, L. (1989). A natural history of negation.

Kaup, B. (2001). Negation and its impact on the accessibility of text information. *Memory & cognition*, *29*(7), 960-967.

Kaup, B. (1997). The processing of negatives during discourse comprehension. In *Proceedings of the nineteenth conference of the Cognitive Science Society* (pp. 370-375).

Kaup, B., Ludtke, J., & Zwaan, R. A. (2005). Effects of negation, truth value, and delay on picture recognition after reading affirmative and negative sentences. In *Proceedings of the annual meeting of the Cognitive Science Society* (Vol. 27, No. 27).

Kaup, B., Lüdtke, J., & Zwaan, R. A. (2006). Processing negated sentences with contradictory predicates: Is a door that is not open mentally closed?. *Journal of Pragmatics*, *38*(7), 1033-1050.

Kaup, B., Yaxley, R. H., Madden, C. J., Zwaan, R. A., & Lüdtke, J. (2007). Experiential simulations of negated text information. *Quarterly journal of experimental psychology*, *60*(7), 976-990.

Kaup, B. & Zwaan, R. A. Effects of negation and situational presence on the accessibility of text information. Journal of Experimental Psychology: Learning, Memory, and Cognition 29, 439–446 (2003).

Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in cognitive sciences*, *4*(12), 463-470.

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

Kutas, M., & Hillyard, S. A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature*, *307*(5947), 161-163.

Lange, K., Kühn, S., & Filevich, E. (2015). " Just Another Tool for Online Studies"(JATOS): An easy solution for setup and management of web servers supporting online studies. *PloS one*, *10*(6), e0130834.

Liuzza, M. T., Candidi, M., & Aglioti, S. M. (2011). Do not resonate with actions: sentence polarity modulates cortico-spinal excitability during action-related sentence reading. *PloS one*, *6*(2), e16855.

Lo, S., & Andrews, S. (2015). To transform or not to transform: Using generalized linear mixed models to analyse reaction time data. *Frontiers in psychology*, *6*, 1171.

Lüdtke, J., Friedrich, C. K., De Filippis, M., & Kaup, B. (2008). Event-related potential correlates of negation in a sentence–picture verification paradigm. *Journal of cognitive neuroscience*, *20*(8), 1355-1370.

Lüdtke, J., & Kaup, B. (2006, July). Context effects when reading negative and affirmative sentences. In *Proceedings of the 28th annual conference of the cognitive science society* (Vol. 27, pp. 1735-1740).

MacDonald, M. C., & Just, M. A. (1989). Changes in activation levels with negation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*(4), 633.

Mannaert, L. N. H., Dijkstra, K., & Zwaan, R. A. (2019). How are mental simulations updated across sentences?. *Memory & Cognition*, *47*(6), 1201-1214.

Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. Behavior research methods, 44(2), 314-324.

Mayo, R., Schul, Y., & Burnstein, E. (2004). "I am not guilty" vs "I am innocent": Successful negation may depend on the schema used for its encoding. *Journal of Experimental Social Psychology*, *40*(4), 433-449.

Nieuwland, M. S., & Kuperberg, G. R. (2008). When the truth is not too hard to handle: An event-related potential study on the pragmatics of negation. *Psychological Science*, *19*(12), 1213-1218.

Nigam, A., Hoffman, J. E., & Simons, R. F. (1992). N400 to semantically anomalous pictures and words. *Journal of cognitive neuroscience*, *4*(1), 15-22.

Nordmeyer, A., & Frank, M. C. (2015). The pragmatics of negation across contexts. In CogSci.

Nordmeyer, A. E., & Frank, M. C. (2014). The role of context in young children's comprehension of negation. *Journal of Memory and Language*, *77*, 25-39.

Orenes, I., Beltrán, D., & Santamaria, C. (2014). How negation is understood: Evidence from the visual world paradigm. *Journal of memory and language*, *74*, 36-45.

Papeo, L., Hochmann, J. R., & Battelli, L. (2016). The default computation of negated meanings. *Journal of cognitive neuroscience*, *28*(12), 1980-1986.

Pea, R. D. (1980). The development ofnegation in early child language. *The socialfoundations oflanguage and thought*, 156-186.

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., ... & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior research methods*, *51*(1), 195-203.

Perea, M., & Rosa, E. (2002). The effects of associative and semantic priming in the lexical decision task. *Psychological research*, *66*(3), 180-194.

Pickering, M. J., & Garrod, S. (2007). Do people use language production to make predictions during comprehension?. *Trends in cognitive sciences*, *11*(3), 105-110.

Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and brain sciences*, *36*(4), 329-347.

Roberts, C. (1996). Information structure in discourse: Toward a unified theory of formal pragmatics. *Ohio State University Working Papers in Linguistics*, *49*, 91-136.

Roberts, C. (2012). Information structure: Towards an integrated formal theory of pragmatics. *Semantics and pragmatics*, *5*, 6-1.

Scappini, M., Delfitto, D., Marzi, C. A., Vespignani, F., & Savazzi, S. (2015). On the non-incremental processing of negation: A pragmatically licensed sentence-picture verification study with Italian (dyslexic) adults. *Cahiers de linguistique française*, *32*, 45-58.

Schneider, W., Eschman, A., and Zuccolotto, A. (2012). E-Prime User's Guide. Pittsburgh: Psychology Software Tools, Inc.

Slobin, D. I. (1966). Grammatical transformations and sentence comprehension in childhood and adulthood. *Journal of verbal learning and verbal behavior*, *5*(3), 219-227.

Song, J. H., & Nakayama, K. (2009). Hidden cognitive states revealed in choice reaching tasks. *Trends in cognitive sciences*, *13*(8), 360-366.

Spivey, M. J., & Dale, R. (2004). On the continuity of mind: toward a dynamical account of cognition.

Spychalska, M., Haase, V., Kontinen, J., & Werning, M. (2019). Processing of affirmation and negation in contexts with unique and multiple alternatives: Evidence from event-related potentials. In *CogSci* (pp. 2845-2851).

Stanfield, R. A., & Zwaan, R. A. (2001). The effect of implied orientation derived from verbal context on picture recognition. *Psychological science*, *12*(2), 153-156.

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, *268*(5217), 1632-1634.

Team, R. (2018). RStudio: integrated development for R. RStudio, Inc, Boston, MA.

Tettamanti, M., Manenti, R., Della Rosa, P. A., Falini, A., Perani, D., Cappa, S. F., & Moro, A. (2008). Negation in the brain: Modulating action representations. *Neuroimage*, *43*(2), 358-367.

Tian, Y. (2014). *Negation processing: A dynamic pragmatic account* (Doctoral dissertation, UCL (University College London)).

Tian, Y., Breheny, R., & Ferguson, H. J. (2010). Why we simulate negated information: A dynamic pragmatic account.

Tian, Y., Ferguson, H., & Breheny, R. (2016). Processing negation without context–why and when we represent the positive argument. Language, Cognition and Neuroscience, 31(5), 683-698.

Tokita, M., & Ishiguchi, A. (2013). Effects of perceptual variables on numerosity comparison in 5–6-year-olds and adults. *Frontiers in Psychology*, *4*, 431.

Tomasino, B., Weiss, P. H., & Fink, G. R. (2010). To move or not to move: imperatives modulate action-related verb processing in the motor system. *Neuroscience*, *169*(1), 246-258.

Trabasso, T., Rollins, H., & Shaughnessy, E. (1971). Storage and verification stages in processing concepts. *Cognitive psychology*, *2*(3), 239-289.

Urbach, T. P., & Kutas, M. (2010). Quantifiers more or less quantify on-line: ERP evidence for partial incremental interpretation. *Journal of Memory and Language*, *63*(2), 158-179.

Vender, M., & Delfitto, D. (2010). Towards a pragmatics of negation: The interpretation of negative sentences in developmental dyslexia. *GG@ G–Generative Grammar at Geneva*, *6*, 1-28.

Wales, R. J., & Grieve, R. (1969). What is so difficult about negation?. *Perception & Psychophysics*, *6*(6), 327-332.

Wason, P. C. (1965). The contexts of plausible denial. Journal of verbal learning and verbal behavior, 4(1), 7-11.

Wason, P. C. (1959). The processing of positive and negative information. *Quarterly Journal of Experimental Psychology*, *11*(2), 92-107.

Wason, P. C. (1961). Response to affirmative and negative binary statements. *British Journal of Psychology*, *52*(2), 133-142.

Wason, P. C., & Jones, S. (1963). Negatives: Denotation and connotation. *British Journal of Psychology*, *54*(4), 299-307.

Whitfield, J. W. (1951). An experiment in psoblem solying. *Quarterly Journal of Experimental Psychology*, *3*(4), 184-197.

Zanellini, S. (2020). The processing of negative sentences: neural and behavioral correlates in healthy volunteers and brain-damaged population.

Zwaan, R. A. (2012). The experiential view of language comprehension: How is negation represented?. Higher level language processes in the brain: Inference and comprehension processes, 255

Zwaan, R. A., & Pecher, D. (2012). Revisiting mental simulation in language comprehension: Six replication attempts. *PloS one*, *7*(12), e51382.

Zwaan, R. A., Stanfield, R. A., & Yaxley, R. H. (2002). Language comprehenders mentally represent the shapes of objects. *Psychological science*, *13*(2), 168-171.