False memories of emotional words as a key for understanding semantic processing in bilinguals and older adults.

Martina Cangelosi

Thesis submitted to obtain the title of
Doctor of Philosophy
At the University of Pavia,
Pavia - Italy.
Spring 2022

Acknowledgments.

I would like to thank my supervisor, Professor Paola Palladino for guiding me in this PhD program, for helping me to define my goals and helping me to reach them. With her positive attitude, she made me feel that I would have been able to be face and overcome any difficulty I could have encountered during this journey.
I would like to thank my other supervisor, Professor Elena Cavallini. I thank her for the pragmatism with which she always advised me. Furthermore, I would like to thank her for her closeness in a difficult moment and for helping me to fit in and feel part of a team that has become a welcoming place that I loved being part of.
I would like to thank Professor Ton Dijkstra, for his precious writing advices, for the generosity and tireless patience with which he dedicated his time to me, encouraging me to improve and, above all, giving me the tools to do so. I also thank him for the humanity with which he has supported and encouraged me, from the beginning of our collaboration and in recent months more than ever.
I would like to thank all my colleagues for the opportunity to share this path, in the best and most complicated moments. Furthermore, I thank them for the coffee breaks, for the shared lunches, for our meetings in the economics department, for the singing moments outside and for the daily train travels together.
I thank the colleagues with whom I collaborated: Francesco Bossi, who patiently helped me to take my first steps in the world of R Studio and OpenSesame, Luca Rinaldi, for helping me to insert a new area into my research and Milvia Cottini for her huge help in data collection and design of the experiment in Bozen.
I thank the undergraduates and trainees who took part in the data collections included in this thesis, for the seriousness and professionalism with which they took part in them.
I thank my parents who are always with me in my choices, who have been part of my path giving me once more their infinite emotional and practical support.
I thank my friends, always present in front of my outbursts, my moments of fatigue or those of despair and participate in the most beautiful ones, for reading endless drafts and for having memorized the name of my doctoral course. I thank my friend Elena for the countless statistical advice and speeches and for the beautiful and enthusiastic look on the world of data analysis that she gave me. I thank my future husband who supported me throughout my PhD course, as he always does, with infinite patience, understanding and love.

Table of contents.

Publications in progress based on this thesis.

- Cangelosi, M., Bossi, F., & Palladino, P. (2021). Did you see that? False memories for emotional words in bilingual children. *Bilingualism: Language and Cognition*, 1–13. https://doi.org/10.1017/s136672892100105x
- Under Review: Cangelosi, M., Barichello, C., Dijsktra, T., & Palladino, P. (2022). To be or not to be bilingual: How reading comprehension and vocabulary size in children are affected by SES. *Reading and Writing Quarterly Journal*.
- In preparation: Cangelosi, M., Cottini, M., & Palladino, P. (2022). The role of emotionality on verbal false memories in German/Italian bilinguals. To be submitted to *Bilingualism: Language and Cognition*.
- In preparation: Cangelosi, M., Rinaldi, L., Dijsktra, T., Palladino, P., & Cavallini, E. (2022). Older adults produce more verbal false memories than younger adults: is it semantics or executive functioning? To be submitted to *Memory and Cognition.*

All the experiments, data collection and analysis are done by me. In all the papers I took the leading role in writing and revising the manuscript. Coauthors provided guidance on various drafts and on revising the manuscripts.

CHAPTER 1

Introduction

If we present people with a list of words, like: "desk, cushion, couch, bench, sit, swivel, sofa, recliner, rocking, sitting" and ask them to remember as many of those words as possible, it is highly probable that they will remember also a word like 'chair' (called "critical lure") as if it were present in the coded list (e.g., Deese, 1959). This phenomenon is called a 'verbal false memory' and it happens in light of the semantic relatedness of the actually coded words with their critical lure.

If we present the same people with yet another list of words, like: "fear, temper, hatred, fury, happy, enrage, emotion, rage, hate, mean", in which the words are associated with the critical lure "anger", they will be more likely to falsely remember the word "anger" than they did with the word "chair" in the previous list. One explanation for this increase in false memories from the first list to the second is that the emotionality of the word "anger" boosts the semantic spreading of activation that causes the verbal false memory for this word relative to 'chair'.

The example shows that false memory depends on the ability to attribute both a deep and a personal meaning to the words to be memorized. As such, the study of verbal false memory is strongly linked to the general cognitive domains of both language and memory. The major aim of the present thesis is to clarify how the connotative component of semantics (i.e., the emotional associations connected with the meaning of a word) acts on false verbal memory in populations that deviate from average (in particular, children, older adults, bilinguals) precisely in these domains.

Verbal False memory and the DRM paradigm.

The existence of false memories may be worrying to some readers. After all, we all place great trust in our memory, entrusting a large part of our conception of reality and what we believe to be true to it. We can accept that we have forgotten about irrelevant details, dates, appointments, or anecdotes. However, never would we be led to believe that our memory can make us remember things or events that never really happened. Yet False Memory occurs when someone believes that an event took place that never really happened.

Why would we develop false memories at all? Having a flexible memory system allows us to imagine the future and to reconsider our past. To complete these cognitive actions, we must have the possibility to recombine episodic and semantic memory elements, producing illusory episodes (*Newman et al., 2009)*.

Verbal false memory research concerns the typology of false memory that has been most extensively studied (e.g., Reyna & Lloyd, 1997). The best known task used to measure it is the Deese, Roediger, and McDermott paradigm (DRM paradigm, Deese, 1959, Roediger & McDermott, 1995). Participants are presented with lists containing words (e.g., "nurse", "hospital", "syringe", "medicine", "gown"…) that are all semantically related to a theme (called "critical lure") not present in the list (e.g., "doctor"). Their task is to remember as many of the coded words as possible. After a delay phase, there is the experimental phase, mostly involving one of two procedural variants. The first variant of the experimental phase, frequently applied, includes a recall phase, in which participants are asked to repeat as many words as possible from the ones memorized during the coding phase without external help. The second variant instead includes a recognition task. In this case, participants are presented with a further list of words, consisting of the actually coded words, while the semantically associated lures are

not presented in the coding phase and while there is a series of distractors that can be semantically related or not related. Participants are asked to recognize and select the actually coded words.

The outcome of both types of experimental phase is that subjects will be likely to recall or recognize the critical but not present lures as already seen or heard. Of crucial importance for both types is that, between memorization and testing, in the delay phase, certain tasks are performed by the participants that intend to distract them from what they memorized just before. Examples of intervening distractor tasks are mathematical tasks, like multiplication or addition.

The first version of the paradigm dates back to 1959 and is due to Deese, who reported one of the most powerful false recall effects ever created in the laboratory. Each of the lists he proposed included 12 words associated with a 'not present' target stimulus, on the basis of available norms of association. Starting from Deese's methodology, the words that compose each list have often been ordered according to their Backward Association Strength (BAS), which is the strength of the association from the coded words to their critical lures (e.g., Arndt, 2011). Deese's work had little impact until 1995, in which year Roediger and McDermott published an article using Deese's lists and extended his results to the recognition version of the paradigm (the involved author list henceforth leading to the abbreviation DRM for the paradigm in question).

To set the stage for the empirical research discussed in the later chapters of this thesis, we will discuss several interesting aspects of Verbal False memory in the following sections. First, we will consider the relationship of Verbal False memory to semantic processing. Subsequently, the effect of emotionality on Verbal False memory is described. In the next part of this Introduction, we will explain the aim and content of each of the empirical

chapters in this thesis. Later, all empirical studies will be integrated into a new theoretical account of the phenomenon of Verbal False memory in the General Discussion concluding this thesis.

Verbal False memory as a result of semantic processing and the DRM paradigm as a key to understand semantic processing (Brainerd *et al.*, 2008).

It has been proven that verbal false memories occur as a consequence of the spreading of semantic activation from the list items to their critical lure (Brainerd, Yang, Reyna, Howe, & Millis, 2008). Although several other theories also account for this mechanism, here we will refer to the two theories that are most strictly related to the focus of the present thesis: Fuzzy Trace Theory (FTT; Reyna et al., 2016) and Spreading Activation Theory (SAT; Roediger *et al.,* 2001).

Fuzzy Trace Theory (FTT; Reyna et al., 2016).

According to this theory, we memorize information on the basis of two different kinds of cognitive processing. On the one hand, *verbatim processing* allows us to represent the surface characteristics of the material we have to memorize. It attends to the material's details and punctual features, such as the position of a word in a DRM list, the color in which it is written, or its length in letters (Brainerd & Reyna, 2005). On the other hand, *gist processing* allows us to understand the general semantic meaning of the information being processed. According to Brainerd and Reyna (2005), we produce false memory on the basis of gist processing. In contrast, a combination of verbatim processing and gist-based processing is mainly responsible for what we correctly remember (accuracy). The two processes, gist and verbatim, proceed in parallel, but their content is retrieved separately. It has been demonstrated that really coded words tend to activate verbatim

details, while critical lures tend to activate gist information (Brainerd & Reyna, 2005). Furthermore, verbatim trace tends to decline more rapidly than gist, which is consistent with the notion that, as time goes by, persons tend to forget the details of an event, but remember its more general meaning (Zhang, 2016).

Spreading Activation Theory (SAT; Roediger *et al.,* 2001).

Roediger et al. (2001) built their Spreading Activation Theory for false memories on the basis of earlier notions (e.g., Anderson, 1983). For them, all words and concepts of a person are stored in a mental lexicon that is semantically organized. Words with similar meanings have stronger associations than weakly related words. Every word corresponds to a node in the mental lexicon that is activated when the word is processed. Activation spreads to the surrounding nodes of semantically-related words and concepts. According to Roediger *et al.,* verbal false memories occur in the DRM paradigm as a consequence of the semantic activation of the actually coded words and their activation spread to the 'not present' semantically-related critical lures.

Brainerd and colleagues (2009) tested the semantic properties of the DRM paradigm. In particular, they analyzed three types of semantic categories. The first category combined seven of Toglia and Battig's (1978) semantic word norms (familiarity, meaningfulness, concreteness, imagery, categorizability, number of attributes, pleasantness). The second one included some emotional dimensions belonging to Bradely and Lang's (1999) word norms (arousal, dominance, valence). The third one included six conceptual relations between pairs words (e.g., target-lure), originally proposed by Wu and Barsalou (2008) (synonymy, antonymy, taxonomy, entity relations, introspective relations, and situational relations). Brainerd and colleagues ran an analysis in which they opposed criterion variables (i.e., false recognition of critical lures, false recall of critical lures, and true

recall of coded words) and predictor variables (i.e., the three groups described above), used to build a semantic profile of the DRM materials. Their results lead to the conclusion that the semantic properties of the DRM materials can account for verbal false memories. More in detail, Brainerd and colleagues argued that, because critical lures score high on semantic dimensions (i.e., categorizability, concreteness, familiarity, imagery, meaningfulness, number of attributes, and pleasantness), any critical prime always activates semantic content to a high degree. Furthermore, they noted that the actually present words are conceptually highly related to their critical lures. Finally, they showed that a higher Backward Association Strength (BAS, see below) does not only correspond with a higher number of false memories (as mentioned in the literature), but also with a higher semantic processing level.

Emotion as a booster for semantic processing.

It has been found that emotional events are better remembered than neutral ones (e.g., Kensinger, 2009). Our emotional reactions are highly adaptive in that they make an event significant and strengthening it as a memory. Interestingly, emotion does not affect only true memory, but also false memory (Zhang, 2016). We can clarify this by considering the nature of emotion in more detail.

Emotionality of words is usually categorized on the two main dimensions of arousal and valence. Arousal refers to the intensity of an emotional stimulus, while valence refers to the negativity or positivity of that stimulus. For example, the word 'shark' has a higher arousal value than 'whale'; and the word 'sunshine' has a positive valence and 'sadness' a negative one.

The main effect of emotionality on false memories comes from arousal. Higher arousal lures are more often falsely remembered that the neutral ones (e.g., Dehon, Larøi, & Van

der Linden, 2010; Sharkawy, Groth, Vetter, Beraldi, & Fast, 2008). This phenomenon occurs because emotional stimuli acquire a sort of 'urgency of elaboration' that results in focusing on the gist of what is being processed instead of on its punctual features. In this sense, one could say that emotionality acts as a booster for semantic processing. Empirical results with respect to the effect of valence on false memory have been less consistent (e.g., Brainerd, Stein, Silveira, Rohenkohol, and Reyna, 2008). Nevertheless, most studies indicate that critical lures with negative valence elicit the largest number of verbal false memories.

It has to be underlined that the role of emotionality on false memories' production seems to be highly sensitive to the nature of the task used or to the way in which emotionality itself is included. For example, Bookbinder and Brainerd (2016) reviewed the literature concerning this aspect and found that when emotionality is included in a task in terms of to-be-remembered events (e.g., words divided according to their emotional features) it boosts the number of false memories, as addressed above. On the contrary, when emotionality is included in a task in terms of mood (e.g., emotional music heard while processing some stimuli) it can protect against false memory. Authors refer to this phenomenon in terms of "context-content paradox": the context refers to the effect played by the mood in which stimuli are processed, while content is the one vehicled by the stimuli themselves. They explain these differences in terms of Fuzzy Trace Theory: in particular, they hypothesize that "variations in emotional content primarily affect memory for the gist of events, whereas variations in emotional context primarily affect memory for events' exact verbatim form". It might be interesting to broaden a context-based view as a future direction, including additional measures aimed at deepening individual emotional traits both in children and in adults and the role they might play on emotional verbal false memories' production. However, for the purposes of the present thesis we

decided to follow a more content-oriented view, using average values obtained through pilot studies we conducted on our target populations (as better explained through the thesis chapters) in order to determine the intrinsical emotional properties of the words. This choice was made on the basis of what we found in the literature (e.g., Brainerd *et al.,* 2008) and in order to avoid burdening an already long experiment.

The Cornell/Cortland Emotional Lists (CEL Lists, Brainerd et al., 2008): an emotional version of the DRM paradigm.

There have been different approaches to creating and administering emotional DRM lists. In general, two main categories of emotional DRM lists are used. One approach divides the lists according to their negative or positive valence. The other approach divides the lists mainly according to their arousal, e.g., anger or sad lists. The first approach has been extensively used in research (e.g., Dehon, Larøi, & Van der Linden, 2010; Brainerd, Stein, Silveira, Rohenkohol, & Reyna, 2008), but it does not allow a clear distinction between valence and arousal effects. Furthermore, in many studies effects of arousal have not been controlled for.

Both valence and arousal are controlled for in the Cornell/Cortland Emotional Lists (CEL Lists, Brainerd *et al., 2008*). This data pool consists of 32 DRM type lists, in which arousal and valence are factorially manipulated. As a result, there are four main categories of DRM lists (and eight lists for every category): high arousal-negative valence, low arousal-negative valence, high arousal-positive valence, and low arousal-positive valence. Mean backward associative strength is controlled for the whole pool. To build the CEL, the authors initially referred to the most commonly used affective norms for English words (ANEW Norms, Bradley, & Lang, 1999) and to the Nelson *et al.*'s norms for word association (1999). The valence and arousal scores are counterbalanced

across the four sets both for the critical lures and for the words composing the lists to be coded. The authors extensively tested their pool on different age populations, from children aged 7 to 11 to adults around 20. In all these cases, they demonstrated that the lists are all effective in eliciting false memories (Brainerd *et al*., 2010).

The procedure according to which they administered their pool of items will be discussed later, in the experimental part of this thesis, since it has been replicated in all the studies that will be presented.

The empirical chapters of this thesis

Having sketched the general theoretical background literature on the phenomenon of verbal false memory and its relationship with the emotionality of the words, we will now explain the subsequent organization of the thesis and give a more detailed presentation of the content of its empirical chapters. The aim and topic of each chapter is specified, followed by a discussion of the participant population and the stimulus materials that will be used for the main task at hand. The thesis will be divided in two main parts, dedicated to research on bilingualism and aging in relation to Verbal False memory.

Part I of the thesis: Bilingualism and Verbal False memory

In the first empirical part of the thesis, chapters 2, 3 and 4, we will analyze how the two dimensions of arousal and valence influence false memories' production in the L2 of different bilingual populations. We chose to start from this topic, since no one considered it before, as far as we reviewed and we think it can be a key to better understand a theme as broad and multifaceted as bilingual word processing is. Delving in the details of Part I, in chapter 2, we discuss the issue for a participant population of 9 to 11 year old children, half of them made up of minority language bilingual children, second generation's immigrants. As a consequence and development of the main project of chapter 2, in

chapter 3 we are going to insert and discuss another study regarding SES in the above mentioned population. In fact, we will show that SES is an important determinant in the categorization of bilingualism that is at least partially able to account for differences between the bilingual populations discerned in this thesis.

In chapter 4, we will replicate this study for another population of 20 to 35 years old adults, half of them consisting of German-Italian bilinguals from Bozen, a fully bilingual reality located in the Northern part of Italy. In this study, we will introduce another variable that turned out to be affected in the DRM paradigm applied to bilinguals, namely reaction time (RT).

Part II of the thesis: Effects of aging and emotion on Verbal False memory

The second part of the thesis will consider the effects of aging and emotionality on Verbal False memory. In particular, in chapter 5, we will discuss the cause of an increased number of verbal false memories reported for older participants (e.g., Devitt & Schacter, 2016; Jacoby & Rhodes, 2006; Norman & Schacter, 1997). In chapter 6, we will briefly consider some preliminary results regarding the role that emotionality plays of verbal false memories in this population. In the following sections, we will motivate our choice for the participant populations in chapter 5 and in chapter 6 we will motivate our manipulation of the emotional aspects of the stimulus lists.

The two participant populations of the thesis

We chose to analyze the populations of bilingual children and older adults, because their performance can be contrasted with respect to the two main dimensions involved in verbal false memory and in semantic processing: language and memory (e.g., Xu, He & Bi, 2017). We focused on bilinguals as a participant population for a number of reasons.

After a long period during which bilinguals were thought to be cognitively disadvantaged, in 1962 Peal and Lambert introduced the opposite idea of a bilingual advantage. When they administered verbal and nonverbal intelligence tests to 10 year old French children and compared them with monolingual peers, they found that bilingual children performed significantly better in both the categories of tests. Later studies deepened these results and found that, in comparison to monolinguals, bilinguals may possess a larger cognitive flexibility (e.g., Green, 1998). Nowadays, a more nuanced view is held, in which bilinguals are not worse or better than monolinguals, but simply different. Bialystok (2008) published an article aiming to smooth sharp theoretical views regarding a possible bilingual advantage. In particular, she distinguished cognitive areas that benefit from being bilingual (i.e., executive functioning, which includes working memory, see Myiake *et al.*, 2000), cognitive areas that draw disadvantages from being bilingual, at least at early stages of literacy (i.e., vocabulary size, as measured in each of the spoken languages), and neutral areas (i.e., long-term memory).

Older adults show a cognitive pattern that seems to be the opposite of the bilingual one. In this context, it has been suggested that healthy aging leads a slow impairment in various cognitive domains. Birren (2006) argued that the forms of memory that are mostly impaired during aging are recollection of information, episodic memory, and working memory. Similarly, the executive functioning domain is strongly affected by aging, as measured by cognitive and neuropsychological tasks (e.g., Fisk & Sharp, 2004; Allain *et al.*, 2005). Vocabulary, on the contrary, seems to be one of the areas that remain intact during aging or even are improved. It has been argued that such areas refer to crystallized knowledge that is more preserved than episodic knowledge (Birren, 2006).

Concerning aging, a possible deteriorization of semantics is still subject of debate (e.g., Birren, 2006). With respect to bilinguals, there are still doubts regarding the role of

semantics relative to language. Is semantics independent from language or is it linked to it? Most theories now argue for a shared semantic system (e.g., Dijkstra, van Heuven, & Grainger, 1998; Dong, Gui & Mcwhinney, 2005), but there is still so much to disambiguate (e.g., Floccia *et al.,* 2020).

With respect to older adults, Birren (2006) reviewed some studies according to which semantic memory should be relatively intact, especially in comparison with episodic memory. Nevertheless, it is hard to determine what effect is a result of semantic processing or due to other nearby but different domains (e.g., executive functioning, considered in more detail in the aging section of this thesis).

The stimulus materials used in the thesis

The CEL Lists adapted to Italian: a pilot study.

In order to adapt and validate the Cornell/Cortland Emotional Lists (CEL, 2008) to Italian, we conducted a pilot study on children, on younger adults and on older adults, comparable in age to that of the studies in this thesis. When the Italian lists were constructed, the frequencies of the words to be included were checked in the CoLFIS Lexical Database (Bertinetto et al., 2005). Arousal and valence were assessed on a dichotomous scale (respectively, high/low and negative/positive). Backward associative strength (BAS) was measured on a 7-point Likert scale (from 1 = not associated at all to 7 = very associated).

Concerning the participant group of children, we administered the Italian version of the CEL materials to 25 monolingual children (M age = 9.8). We compared our Arousal and Valence results to the original ones and found a correspondence of 82% for valence and of 62 % for arousal. We reassigned the lists to the four Arousal * Valence

conditions on the basis of the pilot results and we reordered them on the basis of our BAS results.

For the latter two groups, we recruited 20 healthy older adults (mean age: 74 years) and 20 young adults (mean age: 26.5 years), who did not take part in the actual study. We compared our Arousal and Valence results to the originals and found a correspondence of the 87.4 % in terms of the valence and a correspondence of the 60.6 % for arousal. We again reassigned the lists to the four Arousal * Valence conditions on the basis of the pilot results and reordered them on the basis of our BAS results.

Chapter 7: General Discussion

On the basis of the above summary of background literature and the description of the empirical studies to be reported, the reader will have become aware that the phenomenon of Verbal False memory is investigated in this thesis with respect to many different aspects: participant population characteristics, task demands, stimulus materials, emotional aspects, bilingual status, SES.

All these aspects will be reconsidered in the General Discussion of chapter 7 and integrated within a cohesive and innovative theoretical view of Verbal False memory.

As a final remark, we note that there is a lot more to be said about Verbal False memory. However, to avoid reduplication, the detailed description of further relevant aspects has been reserved for the introduction of each of the following empirical chapters. Nevertheless, a limited amount of redundancy across the thesis remains unavoidable, because some of the chapters consist of papers already published or under review. It is our hope that the remaining degree of redundancy will be beneficial to the reader, in the sense that each individual chapter of the thesis can now be read on its own.

References

Allain, P., Nicoleau, S., Pinon, K., Etcharry-Bouyx, F., Barré, J., Berrut, G., ... & Le Gall, D. (2005). Executive functioning in normal aging: A study of action planning using the Zoo Map Test. *Brain and cognition, 57(*1), 4-7.

Anderson, J. R. (1983). A spreading activation theory of memory. *Journal of Verbal Learning and Verbal Behavior, 22*, 261–295.

Bertinetto P. M., Burani C., Laudanna A., Marconi L., Ratti D., Rolando C., Thornton A. M. (2005). Corpus e Lessico di Frequenza dell'Italiano Scritto (CoLFIS).

Bialystok, E. (2009). Bilingualism: The good, the bad, and the indifferent. *Bilingualism: Language and cognition, 12*(1), 3-11.

Birren, J. E. *Handbook of the Psychology of Aging, Sixth Edition (Handbooks of Aging) (2005–12-23*). (2021). Academic Press; 6 edition (2005–12-23).

Bookbinder, S. H., & Brainerd, C. J. (2016). Emotion and false memory: The context–content paradox. *Psychological Bulletin, 142*(12), 1315–1351. https://doi.org/10.1037/bul0000077

Bradley, M. M., & Lang, P. J. (1999). *Affective norms for English words (ANEW): Instruction manual and affective ratings* (Vol. 30, No. 1, pp. 25-36). Technical report C-1, the center for research in psychophysiology, University of Florida.

Brainerd, C. J., & Reyna, V. F. (2005). *The Science of False Memory (Oxford Psychology Series, 38)* (1st ed.). Oxford University Press.

Brainerd, C., Stein, L., Silveira, R., Rohenkohl, G., & Reyna, V. (2008). How Does Negative Emotion Cause False Memories? *Psychological Science, 19* (9), 919–925. https://doi.org/10.1111/j.1467-9280.2008.02177.x

Brainerd C. J., Yang Y, Toglia MP, Reyna VF, Stahl C. (2008). Emotion and false memory: The Cornell/Cortland norms. *In annual meeting of the Psychonomic Society, Chicago, IL.*

Brainerd, C. J., Yang, Y., Reyna, V. F., Howe, M. L., & Mills, B. A. (2008). Semantic processing in "associative" false memory. *Psychonomic Bulletin & Review, 15*(6), 1035–1053. https://doi.org/10.3758/pbr.15.6.1035

Brainerd, C., Holliday, R., Reyna, V., Yang, Y., & Toglia, M. (2010). Developmental reversals in false memory: Effects of emotional valence and arousal. *Journal of Experimental Child Psychology, 107* (2), 137–154. https://doi.org/10.1016/j.jecp.2010.04.013

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology, 58*(1), 17–22. https://doi.org/10.1037/h0046671

Devitt, A. L., Tippett, L., Schacter, D. L., & Addis, D. R. (2016). Autobiographical memory conjunction errors in younger and older adults: Evidence for a role of inhibitory ability. *Psychology and Aging, 31*(8), 927–942. https://doi.org/10.1037/pag0000129

Dehon, H., Larøi, F., & van der Linden, M. (2010). Affective valence influences participant's susceptibility to false memories and illusory recollection. *Emotion, 10* (5), 627–639. https://doi.org/10.1037/a0019595

Dong, Y., Gui, S., & Macwhinney, B. (2005). Shared and separate meanings in the bilingual mental lexicon. *Bilingualism: Language and Cognition, 8*, 221–238. DOI: 10.1017/s1366728905002270

Fisk, J. E., & Sharp, C. A. (2004). Age-related impairment in executive functioning: Updating, inhibition, shifting, and access. *Journal of clinical and experimental neuropsychology, 26* (7), 874-890.

Floccia, C., Delle Luche, C., Lepadatu, I., Chow, J., Ratnage, P., & Plunkett, K. (2020). Translation equivalent and cross-language semantic priming in bilingual toddlers. *Journal of Memory and Language, 112*, 104086.

ISO 690

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and cognition, 1* (2), 67-81.

van Heuven, W. J. B., Dijkstra, T., & Grainger, J. (1998). Orthographic neighborhood effects in bilingual word recognition. *Journal of Memory and Language, 39*, 458–483. DOI: 10.1006/jmla.1998.2584

Jacoby, L. L., & Rhodes, M. G. (2006). False Remembering in the Aged. *Current Directions in Psychological Science, 15* (2), 49–53. https://doi.org/10.1111/j.0963-7214.2006.00405.x

Kensinger, E. A. (2009). Remembering the details: Effects of emotion. *Emotion review, 1* (2), 99-113.

Miyake, A., Emerson, M. J., & Friedman, N. P. (2000). Assessment of executive functions in clinical settings: Problems and recommendations. In *Seminars in speech and language (Vol. 21, No. 02, pp. 0169-0183).* Copyright© 2000 by Thieme Medical Publishers, Inc., 333 Seventh Avenue, New York, NY 10001, USA. Tel.:+ 1 (212) 584-4663.

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (1999). The University of South Florida word association, rhyme, and word fragment norms. *Unpublished manuscript, University of South Florida, Tampa*. Neuschatz,

Newman, E. J., & Lindsay, D. S. (2009). False memories: What the hell are they for? *Applied Cognitive Psychology, 23* (8), 1105–1121. https://doi.org/10.1002/acp.1613

Norman, K. A., & Schacter, D. L. (1997). False recognition in younger and older adults: Exploring the characteristics of illusory memories. *Memory & Cognition, 25*(6), 838–848. https://doi.org/10.3758/BF03211328

Peal, E., & Lambert, W. E. (1962). The relation of bilingualism to intelligence. *Psychological Monographs: general and applied, 76* (27), 1.

Reyna, V. F., & Lloyd, F. (1997). Theories of false memory in children and adults. *Learning and Individual Differences, 9* (2), 95–123. https://doi.org/10.1016/s1041-6080(97)90002-9

Reyna, V. F., Corbin, J. C., Weldon, R. B., & Brainerd, C. J. (2016). How fuzzy-trace theory predicts true and false memories for words, sentences, and narratives. *Journal of Applied Research in Memory and Cognition, 5* (1), 1–9. https://doi.org/10.1016/j.jarmac.2015.12.003

Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21* (4), 803–814. https://doi.org/10.1037/0278-7393.21.4.803

Roediger III, H. L., Balota, D. A., & Watson, J. M. (2001). Spreading activation and arousal of false memories.

Sharkawy, J. E., Groth, K., Vetter, C., Beraldi, A., & Fast, K. (2008). False Memories of Emotional and Neutral Words. *Behavioural Neurology, 19* (1–2), 7–11. https://doi.org/10.1155/2008/587239

Toglia, M. P., & Battig, W. F. (1978). Handbook of semantic word norms. Lawrence Erlbaum

Wu, L. L., & Barsalou, L. W. (2009). Perceptual simulation in conceptual combination: Evidence from property generation. *Acta psychologica, 132* (2), 173-189.

Xu, Y., He, Y., & Bi, Y. (2017). A Tri-network Model of Human Semantic Processing. *Frontiers in Psychology, 8*. https://doi.org/10.3389/fpsyg.2017.01538

Zhang, W., Gross, J., & Hayne, H. (2016). The effect of mood on false memory for emotional DRM word lists. *Cognition and Emotion, 31*(3), 526–537. https://doi.org/10.1080/02699931.2016.1138930

Section 1

The influence of emotion on verbal false memories in bilinguals

Theoretical issue

Here we will discuss what is going to happen in chapter 2, 3 and 4 about the effect of emotion on verbal false memories in bilinguals. Now we will set the stages by discussing the literature background and anticipating the materials and experimental designs. As explained in the Introduction, a verbal false memory occurs when people falsely remember a word (called "critical lure") that was not present in just seen list of words that are all semantically related (e.g., Deese, 1959; Roediger & McDermott, 1995). This phenomenon has been studied extensively in monolinguals, but when I did a preliminary review of the literature, we found almost an absence of studies on how the emotional component of the lexicon affects the production of verbal false memories in bilinguals.

In monolinguals, the phenomenon of verbal false memory has been argued to arise as a consequence semantic activation spreading during word coding (Brainerd, 2008). The connotative dimension of semantics (i.e., the emotionality conveyed by the meaning of the items) has been found to play a significant role in false memory production. Indeed, a higher degree of emotionality of the items in the test list elicits a larger number of false memories (e.g., Zhang, 2016).

We set out to study this phenomenon and the role played by emotion in bilingual false memory, for word processing in the bilinguals' second language (L2). We wished to test two possible underlying mechanisms. On the one hand, bilinguals tend to process word meaning as part of a shared system between the two languages they speak (e.g., van Heuven, Dijkstra, & Grainger, 1998; Dong, Gui, & Mcwhinney, 2005). This leads to the hypothesis that bilinguals might produce as many verbal false memories as monolinguals do. On the other hand, there is contrasting evidence about the emotionality of words in L2. Although some studies indicated that the emotionality of the L2 and L1 lexicons are comparable (e.g., Altarriba & Canary, 2004), many more found that the L2 lexicon may

be less emotionally connotated than the bilinguals' L1 lexicon or the monolingual's lexicon (e.g., Degner, Doycheva, &Wentura, 2012). On this basis, one might hypothesize that bilinguals are less influenced by emotionality in the production of verbal false memories than monolinguals.

Participant profile

Two different populations of bilinguals were tested. One group of participants was composed of children of second-generation immigrants speaking a minority language as L1. These children speak different languages that are not included in the cultural background of the society in which they live, namely Italy (more details about the languages will be given in chapter 2). They learnt Italian at school and they are proficient in it. On average, they have a lower socioeconomic status than their monolinguals peers and a lower linguistic status for their mother tongue, since their social context does not promote a strong affiliation with their L1.

The other group of participants consisted of adults living in Bozen. They are German/Italian bilinguals. On average, they have an high socioeconomic status. The social context in which they live highly promotes bilingualism at every level of their society (i.e., bilingual schools and fully bilingual social context).

What prompted us to focus on these two different populations is the awareness that their characteristics make them two completely and opposite types of bilinguals. Contrasting their performance in research could clarify commonalities and distinctions between them and their processing. Indeed, the hypothesis is that both the semantic and emotional pattern processing involved in verbal false memories might be different between them. In particular, our prediction was that minority language bilingual children

(30 % of the sample speaking Arabic; 10 % speaking Chinese, 27 % speaking South American Spanish, 7 % speaking Albanian, 5 % speaking Romanian, 3 % speaking Polish, 4 % speaking Brazilian Portuguese and 1 % speaking Turkish and 13 % speaking other dialects) could have been so immersed in a context that promotes their L2 (Italian) at the expense of their L1, so as to perform similar to their monolingual (Italian) peers in the semantic and emotional patterns, especially considering their age and school level (9-11 years).

Our prediction for the second population, that of German/Italian bilinguals, was that the semantic pattern causing verbal false memories would be similar to that of their monolingual peers. However, at the same time, we expected them to be less influenced by the emotional component of their L2 lexicon.

The role of socioeconomic status (SES)

The two described populations are connected and contrasted in terms of their Socioeconomic status (SES). Socioeconomic status is also the factor that led us to hypothesize that the two populations might show different verbal false memories patterns. Morton and Harper (2007) were among the first researchers to argue that taking socioeconomic status into account when studying the performance of bilinguals is crucial. In fact, they pointed out that it could be misleading to make inferences about a possible bilingual advantage without considering this dimension of SES. More recently, Naeem, Filippi, Periche-Tomas, Papageorgiou, and Bright (2008) came to the same conclusion when they were addressing the decisive modulating role of SES. In order to measure SES, we administered the Hollingshead Four Factor Index of Social Status Questionnaire

(1975), in its original version with respect to the children group and in an adapted version with respect to adults' group.

Task design

The DRM task design that will be adopted will be identical for the two bilingual populations (first study of chapters 2 and 3). Bilingual participants will be presented with a version of the Cornel/Cortland Emotion Lists (CEL Lists; Brainerd *et al.*, 2008) adapted for their L2, Italian. Each participant group will be compared with an Italian monolingual control group of the same age, who will have the same DRM materials administered, but in their L1 (i.e., Italian).

Control variables

In both studies, we included several control variables. As control variables common to the two populations, we considered short-time memory, vocabulary size, and categorical fluency. However, in response to specific dimensions that have to be considered for each group, different control variables were also included. For the monolingual children group, we added a measure of reading comprehension, because it is an important determinant for children of this age group and when dealing with linguistic variables (e.g., Bonifacci, Lombardo, Pedrazzini, Terracina, & Palladino, 2020). For the German/Italian bilingual adults, we also examined the reaction times obtained in the DRM task.

Finally, we administered a questionnaire to measure linguistic background. In particular, children were presented with the adapted to Italian version of the Paradis's ALDeQ questionnaire developed by Bonifacci, Mari, and Porrelli (2016). Adults were presented with a questionnaire that was built *ad hoc* by Cottini and Basso (2014).

An additional study

In addition to the two main studies described in chapters 2 and chapter 4, in chapter 3 we chose to insert a third one derived from our study on minority language bilingual children on the role of SES in relation to variables like vocabulary and comprehension (Cangelosi, Bossi, & Palladino, 2021). This choice was motivated by the demonstrated importance of SES for this bilingual population, its coherence with the previously reported SES literature, and previous studies addressing the importance of this variable for linguistical skills in a developmental phase (e.g., Bonifacci *et al.*, 2020).

References

Altarriba, J., & Canary, T. M. (2004). The Influence of Emotional Arousal on Affective Priming in Monolingual and Bilingual Speakers. *Journal of Multilingual and Multicultural Development, 25*(2–3), 248–265. https://doi.org/10.1080/01434630408666531

Bonifacci, P., Lombardo, G., Pedrinazzi, J., Terracina, F., & Palladino, P. (2019). Literacy Skills in Bilinguals and Monolinguals with Different SES. *Reading & Writing Quarterly, 36* (3), 243–259. https://doi.org/10.1080/10573569.2019.1635057

Bonifacci, P., Mari, R., Gabbianelli, L., Ferraguti, E., Montanari, F., Burani, F., & Porrelli, M. (2016). Sequential bilingualism and specific language impairment: The Italian version of ALDeQ parental questionnaire. *BPA-Applied Psychology Bulletin* (Bollettino di Psicologia Applicata), *64*, 50–64.

Brainerd, C. J., Yang, Y., Reyna, V. F., Howe, M. L., & Mills, B. A. (2008). Semantic processing in "associative" false memory. *Psychonomic Bulletin & Review, 15*(6), 1035–1053. https://doi.org/10.3758/pbr.15.6.1035

Brainerd C. J., Yang Y, Toglia MP, Reyna VF, Stahl C. (2008). Emotion and false memory: The Cornell/Cortland norms. In *annual meeting of the Psychonomic Society, Chicago, IL*.

Poster presentation: Cottini, M. & Basso, D. Prospective memory development in monolingual and bilingual children. *IV International Conference on Prospective Memory (ICPM4),* Naples, 26-30 May 2014.

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology, 58*(1), 17–22. https://doi.org/10.1037/h0046671

Degner, J., Doycheva, C., & Wentura, D. (2012). It matters how much you talk: On the automaticity of affective connotations of first and second language words. *Bilingualism: Language and Cognition, 15* (1), 181-189.

Dong, Y., Gui, S., & Macwhinney, B. (2005). Shared and separate meanings in the bilingual mental lexicon. *Bilingualism: Language and Cognition, 8*, 221–238. DOI: 10.1017/s1366728905002270

van Heuven, W. J. B., Dijkstra, T., & Grainger, J. (1998). Orthographic neighborhood effects in bilingual word recognition. *Journal of Memory and Language, 39*, 458–483. DOI: 10.1006/jmla.1998.2584

Morton, J. B., & Harper, S. N. (2007). What did Simon say? Revisiting the bilingual advantage. *Developmental Science, 10* (6), 719–726. https://doi.org/10.1111/j.1467-7687.2007.00623.x

Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21* (4), 803–814. https://doi.org/10.1037/0278-7393.21.4.803

Zhang, W., Gross, J., & Hayne, H. (2016). The effect of mood on false memory for emotional DRM word lists. *Cognition and Emotion, 31*(3), 526–537. https://doi.org/10.1080/02699931.2016.1138930

CHAPTER 2

Did you see that? False memories for emotional words in bilingual children

Cangelosi, M., Bossi, F., & Palladino, P. (2021). *Bilingualism: Language and Cognition*, 1–13. https://doi.org/10.1017/s136672892100105x

Abstract

When participants process a list of semantically strongly related words, the ones that were not presented may later be said, falsely, to have been on the list. This 'false memory effect' has been investigated by means of the DRM paradigm. We applied an emotional version of it to assess the false memory effect for emotional words in bilingual children with a minority language as L1 (their mother tongue) and a monolingual control group. We found that the higher emotionality of the words enhances memory distortion for both the bilingual and the monolingual children, in spite of the disadvantage related to vocabulary skills and of the socioeconomic status that acts on semantic processing independently from the condition of bilingualism. We conclude that bilingual children develop their semantic knowledge separately from their vocabulary skills and parallel to their monolingual peers, with a comparable role played by Arousal and Valence.

Introduction

Mariana is ten years old and she attends the sixth grade in Genoa. Her family moved to Italy, to Mariana's father's hometown when she was just three years old, leaving Santiago del Chile, Mariana's mother's birthplace. The latter speaks to Mariana in Spanish, to maintain a bond with the child's mother tongue. Mariana has always been frightened by the idea of the hospital, by the doctors, with their white coats and serious faces, by the smell of disinfectant that invades the rooms, by the beds, the comings and goings of people, the syringes. The words doctor, sterile gown, cot, syringe have a priority meaning for Mariana in light of her fear.

The emotional component of the lexicon influences its processing. During the course of the present study we will examine how the emotional word processing develops in the second language of proficient bilingual children and how it acts on their semantic elaboration: does the emotionality of the lexicon play the same role in one's second language, making semantic based processing a priority, as it does in one's first language?

False memory

One way to pursue our aim is to measure VERBAL FALSE MEMORIES (Oliveira, Albuquerque & Saraiva, 2018; Seamon, Luo, Schlegel, Greene & Goldenberg, 2000). Suppose we presented Mariana with a list of words, all semantically related to the concept of hospital (e.g., "doctor, sterile gown, bed, syringe"), without ever mentioning the word "hospital" (often called CRITICAL LURE). Suppose further that we then asked her for a list of words she remembers of the words we presented her with. It is likely that she would also mention the word "hospital" – the lure, activated by those actually seen, not present but related at the level of meaning.

To give a definition of this phenomenon, VERBAL FALSE MEMORY occurs when participants "claim that certain words that have not actually been presented are as distinct as those that actually appear as equivalents" (Zhou, 2019).

The question could then be: why does this kind of memory distortion happen? The way in which our memory system works is adaptive and false memories fall within that system (Schacter, Guerin & St. Jacques, 2011), leading to an extensive and articulated understanding of events that may not be well-coded at first (Schacter, 2012); having a flexible memory system enables us to imagine the future and to reconsider one's past (Klein, Robertson & Delton, 2009). To complete these cognitive actions we need to have the possibility to recombine episodic and semantic memory elements. This recombination can produce distortions in remembrance, at the expense of the accuracy and in favour of a more global vision (Newman & Lindsay, 2009).

False memory in children

False memory has been studied not only in an adult population, but also over the course of life and, in particular, during childhood (e.g., Reyna & Lloyd, 1997; Bjorklund, 2000; Brainerd, Reyna & Forrest, 2002; Metzger, Warren, Shelton, Price, Reed & Williams, 2008; Otgaar & Smeets, 2010; Otgaar, Howe, Muris & Merckelbach, 2019). The analysis of this phenomena in a developmental phase, made it possible to clarify its onset and its relationship with the performance in other aspects of a cognitive system that is still in formation (e.g., Brainerd, Reyna & Holliday, 2018). For a long time it was believed that younger children were more prone to the formation of false memories: the cause of this has been traced back to their lower ability to memorize specific details of events related to their lives (e.g., Binet, 1900; Bruck & Ceci, 1999; Quas, Qin, Schaaf & Goodman,

1997; Reyna, Mills, Estrada & Brainerd, 2007; Small, 1896; Stern, 1910). However, Brainerd *et al.* (2002) found that false memories increase proportionally with age while recalling or recognizing verbal material. The authors refer to this phenomenon in terms of "developmental reversals" (Brainerd, Holliday, Reyna, Yang & Toglia*,* 2010) and attribute it to the fact that the more children grow, the more they acquire the ability to "get the gist" of a series of items consistent in their semantics.

Metzger *et al.* (2008) conducted three experiments using word lists to analyse developmental trends in accurate and false memory production for words. At first, they adapted their materials to be consistent with 2nd and 8th grader children's vocabulary, then they asked participants to generate their own lists. The authors found that both true and false memory increased with age. Furthermore, they found that young children did have access to their semantic storage, but they didn't seem to be able to use it to improve their recall. Their results lead to support Brainerd *et al.* (2010) thesis of the developmental reversals, by affirming that the more children grow, the larger use they make of their semantic knowledge, which leads to better recall the actually presented words, but also to a larger semantic-based misremembrance.

False memory in bilinguals: A key to access bilingual word form and meaning representations in adults and children

The modality of access to language in bilinguals is affected by various features (e.g., proficiency, age of acquisition and language dominance, e.g., Kastenbaum, Bedore, Peña, Sheng, Mavis, Sebastian-Vaytadden, Rangamani, Vallila-Rohter & Kiran, 2019; Montrul & Foote, 2012; Puig-Mayenco, Cunnings, Bayram, Miller, Tubau & Rothman, 2018) and it is still the subject of a wide debate. False memory can provide insights related to semantic processing in bilinguals and to the way in which they process different

categories of stimuli, in comparison with monolinguals. The ease with which a word can be retrieved from memory depends on the goodness of fit between the signal and the stored representation, the memory strength of a word, and the number of words that partially match the speech signal and, as a result, compete for selection with the target word. This process is effortless for monolingual speakers, but may be more challenging for second language (L2) and bilingual speakers (Schmidtke, 2014). The analysis of this effort can shed light on some currently active debates regarding bilingual word form and meaning representations, both when we consider adults and children.

A still open question regards indeed how bilinguals represent and store meaning in each of their languages (Dong, Gui & Macwhinney, 2005; Altarriba & Basnight-Brown, 2009; Storms, Ameel & Malt, 2015; Braunstein, Ischebeck, Brunner, Grabner, Stamenov & Neuper, 2012; Khateb, Pegna, Michel, Mouthon & Annoni, 2016; Francis, 2018; Novitskiy, Myachykov & Shtyrov, 2018 and Floccia, Delle Luche, Lepadatu, Chow, Ratnage & Plunkett, 2020): one of the key aspects is whether bilinguals access a shared or separate semantical store. The answer can explain much about language-dependency of semantics.

van Heuven, Dijkstra and Grainger (1998) started from the theoretical framework of an implemented version of the Bilingual Interactive Activation (BIA) model to account for a series of results, obtained through progressive demasking and lexical decision tasks, in order to examine how Dutch/English bilinguals recognized target words belonging to one language and how this processing was affected by orthographic neighbourhood from the same language used in the task or from the other language of participants. Their results provided evidence for parallel activation of words in an integrated Dutch/English lexicon.

Dong *et al.* (2005) proposed that the mental lexicon in bilinguals was represented through a shared and distributed model. They conducted two experiments to demonstrate their view. The first one was a semantic priming paradigm and led to the demonstration of the sharing of relations between concepts across translation equivalents. The second aimed at examining the details of meaning separation, and its results showed that bilinguals have the tendency to integrate meanings through translation equivalents, but the L1 words are preferably represented in the L1 conceptual system, while the L2 words tend to be maintained in the L2 conceptual one.

The organization of the mental lexicon in bilinguals and the way they store meanings has been extensively studied also on children. Kajcsa (2010) assessed it on a bilingual sample of children attending kindergarten through a picture-naming task. Children of the age of 3-4 years appear to be able to access their mental lexicon in a sophisticated way and they show the same pattern found in an adult population, with the only different tendency for children to preferentially process nouns. Furthermore, bilingual children rely on the mental lexicon of both languages, linked to the same conceptual system.

Although today most literature seems to propend for the idea of an integrated semantic system for bilinguals' representation of meanings, this topic still involves many researchers (e.g., Gathercole, 2020).

We think that false memory can help better understanding the above mentioned aspects. Graves and Altarriba (2014) reviewed some of the main results related to false memories for verbal material in bilinguals. Anastasi, Rhodes, Marquez and Velino (2005) tested false memories for words in monolingual English and bilingual participants and found that the bilinguals showed a reduction in false alarms in comparison with the English monolinguals for English words, but did not show a significant false memory

reduction for their native language; Kawasaki-Miyaji *et al.* found that a larger amount of false recognition occurred in Japanese (the participants' dominant language) than in English, but in both the languages false memories were produced (Graves & Altarriba, 2014). In general, it was found that false memories occurred in the between-language conditions and this result is consistent with views that assume the existence of separate lexical representations for a word in each language, but a common conceptual-level representation for both languages.

Marmolejo, Diliberto-Macaluso and Altarriba (2009) presented Spanish-English bilinguals with 10 lists of words, aiming to examine the impact of switching language on false memory. Results showed a higher proportion of veridical and false recall in English, the dominant L2 language, than in Spanish, which was the participants' L1.

Howe, Gagnon and Thouas (2008) examined false memories for words in bilingual children, within and between languages through a recall task and through a recognition task; the children were aged from 6 to 11 years old. The authors found an increasing production of false memories as age increased, coherently with the previously exposed hypothesis of a larger *gist* elaboration during growth, as found in monolinguals. Furthermore, children of all the age groups exhibited a better recall of truly presented words in the within-language condition; as regards the false recall, younger children showed fewer false memories in the between-language condition, while the older children and the adults showed the opposite pattern. To conclude, as regards the recognition task, all the age groups showed a larger number of false memories in the between language condition.

As Riesthuis, Otgaar and Wang concluded (2019), the results found in the literature regarding false memories in bilinguals' L2 is rather controversial and there is

still need to assess to what extent the false and true memories are encoded in a similar way by bilinguals and by monolinguals.

Deese-Roediger-McDermott (DRM) Paradigm

The DRM paradigm is the task that is most commonly used to measure false memory. Its name is an acronym of its three authors: it was developed by Deese at first (1959) and then implemented by Roediger and McDermott (1995). Its most famous version consists in lists of words, all semantically related to a lure that represents the semantic nucleus or theme of each list and that won't be presented. The task is divided in two different moments: at first, participants are exposed to a coding phase, during which they have to memorize as many words as possible from the lists. Secondly, usually after a period of latency, there is the experimental phase, when they have to remember the words they have previously memorized. There are two variants of the experimental phase: subjects can be required to simply recall the words they remember, or they may have to recognize them among other not presented words, among which there are also the critical lures related to every presented list (i.e., word recognition, defined as a "complex process that requires the encoding of a signal and subsequent mapping of this information to representations in memory", Schmidtke, 2014). In both these two variants of the task people are supposed to falsely recall or recognize also the lures, on the basis of the spread of semantic activation that arises when the words are coded and that reinforces the overdetermined concept that ties them.

Brainerd *et al.* (2008b) assessed the semantic properties of the DRM lists, through the examination of 16 dimensions (i.e., familiarity, meaningfulness, concreteness, imagery, categorizability, number of attributes, pleasantness, arousal, dominance, valence, synonymy, antonymy, taxonomy, entity relations, introspective relations, and

situational relations) and found through a factorial analysis that they are a valid index of semantics, extremely rich in meaning.

False memory as a semantic index

On the one hand, as we already mentioned, false memory can be meant as an inaccuracy and a sign of the fallibility of our mnestic system. However, on the other hand, it has been proved that it arises as a result of a semantic-based processing of the information.

Semantics is the study of meanings in natural or artificial languages. In psychology it concerns «how human users of language come to be able to understand what utterances in a language mean. Necessarily, it is also concerned with the question of how the meanings of words are represented in the mind» (Sanford, 2006). Semantics mediates our ability to understand the relationship between things, as well as to analyze and categorize the world around us, giving our knowledge an order; the desire to better understand it has moved a large part of research and has opened an interdisciplinary debate, which involves various fields, from psychology, to philosophy, linguistics and others.

Fuzzy Trace Theory (Reyna & Brainerd, 1995) accounts for false memory as a semantic index, postulating that recollection occurs on the basis of two different and independent forms of processing: *verbatim* retrieval concerns the elaboration of the surface characteristics of a trace to be memorized (e.g., the colour of a stimulus, its position among other elements); conversely, *gist* retrieval refers to the extrapolation of the meaning conveyed by what is being processed, namely its semantics. Both the verbatim and the gist form of elaboration act in the formation of false memory: while the former is deputed to select the actually presented material and reject the critical lures, the latter leads to processing the information on the basis of coherence and semantic

similarity, and it acts in the unfolding of false memory, by promoting acceptance of the non-present but associated critical lures. These two ways act oppositely, but complementarily, respectively explaining the level of accuracy and the level of false memory (Reyna, Corbin, Weldon & Brainerd, 2016).

Another theory that accounts for false memory production in terms of semantic processing is the associative activation theory (AAT; Howe, Wimmer, Gagnon & Plumpton, 2009): it rises from the spreading activation models and it postulates that the processing of one concept makes the activation to the related concept nodes spreading in one's knowledge. This activation leads to the awakening of other related concepts, and the "incorrectly" activated ones may lead to be later falsely remembered as already seen or heard and, though it may account for false memory production. The stronger the relations between concepts and the more one's knowledge base spreads fast, the more false memory will be likely to occur (Otgaar, Howe, Muris & Merckelbach, 2018).

Emotion boosts semantic processing

Both memory and the elements to be memorized are never totally aseptic. On the contrary, they are inseparable from the emotional component and imbued with it. This aspect is part of the very concept of semantics and in linguistics it is called "connotation": with this term we refer to the profound meaning of a term, imbued with beliefs and emotions (Riemer, 2010). Zhang, Gross and Hayne (2016) underlined that emotion is one of the factors that acts on the quality of our memories: the higher emotional connotation of a word represents a preferential channel in terms of word processing and it is able to boost the semantic activation. It has been proved that the degree of emotionality of words acts not only on true recollection, but also on memory distortion, being able to elicit a larger number of false memories (Kaplan, Van Damme, Levine & Loftus, 2015).

This fact goes in the direction of an enhanced gist-based processing for emotional words, which responds to a greater activation and urgency, the same found outside the laboratory, in a context of "everyday life", in which the “emotional information receives pr2eferential processing, which facilitates adaptive strategies for survival” (Lee, Greening & Mather, 2015).

What primarily acts on memory is the arousal of the elicited emotions: indeed, from a phylogenetic point of view, a high arousal event is more likely to be related to personal survival factors or reproductive success, compared to neutral stimuli and it has a highly adaptive value.

Not only the arousal, but also the valence associated to a stimulus acts on its recollection. In particular, it seems that experiences with a negative valence are remembered with a greater number of details compared to neutral and positive valence memories (Zhang *et al.*, 2016).

Emotional words processing in bilinguals.

Memorization of emotive words in bilinguals has been assessed at first by Anooshian and Hertel (1994), who compared recall of emotional and neutral words through a rating task, in either a first and a second language. These two authors hypothesized that the emotional words in a second language lack the emotional connotation and, consequently, are better recalled, acting similarly to the neutral ones. They tested a group of Spanish – English bilinguals and confirmed their hypothesis of a less emotionally charged lexicon in L2.

Ayçiçeği and Harris (2004) started from the methodology of Anooshian and Hertel to demonstrate that, on the contrary, emotion plays a role also in the bilingual L2. In particular, the authors revised the materials used by Anooshian and Hertel noticing that they only considered the arousal of the words (emotional vs neutral). Ayçiçeği and Harris

implemented their set of emotional words, considering valence (positive vs negative), arousal (emotional versus neutral). From their results, obtained by administering a recall and a recognition task to native Turkish bilinguals, speaking English as L2, it appears that emotion is able to elicits an enhancement of memorization. They found this same patten both in L1 and in L2; in particular, they found that negative words were better remembered than positive ones, supporting the proposal that bilingual speakers elaborate negative stimuli better than positive ones in their L2, that led to smaller differences between the two languages.

A similar result has been confirmed by Ferré, García, Fraga, Sánchez-Casas and Molero (2010), who administered a set of positive, negative and neutral stimuli to different groups of bilinguals. From their study it appeared that emotional words were better recalled than neutral ones and that valence and arousal act in interaction, playing a mutual influence on memory results. This pattern was found in the bilinguals' first and second language and regardless of language dominance, context of acquisition, age of acquisition and similarity between languages. In a further and similar study conducted in 2017, they demonstrated that language status plays an important role in emotional word processing. Also Chen, Lin, Chen, Lu and Guo (2015) examined the elaboration of emotional lexicon in both the languages spoken by Chinese-English bilinguals and found that in L1 it would be processed more rapidly than the neutral one, while in L2 the emotional word processing would rely more on semantics. Brase and Mani (2017) assessed the effects of learning context on the acquisition and processing of emotional words in bilinguals. From their results, they found that the context of acquisition varied depending on the language status of the L1 versus the L2 and on the requirement of the tasks. Their results have been replicated by Grabovac and Pléh (2014).

Emotional DRM lists

The DRM paradigm allows us to understand the influence that the emotional content of the lexicon has on its memorization, in both an adult and in a child population (e.g., Brainerd, Yang, Reyna, Howe & Mills, 2008b; Brainerd, Yang, Toglia, Reyna & Stahl, 2008c; Howe, 2007; Quas, Rush, Yim, Edelstein, Otgaar & Smeets, 2015; Baugerud, Howe, Magnussen & Melinder, 2016). Brainerd *et al.* (2008c) found that previously used DRM lists differed in both the arousal and in valence dimensions. On the basis of this finding, the authors validated the Cornell/Cortland Emotional Lists (CEL, Brainerd *et al.*, 2008c). This is a pool of 32 lists, divided into 4 subsets: high arousal and negative valence lists, low arousal and negative valence lists, high arousal and positive valence lists, low arousal and positive valence lists. The administration of lists belonging to each of the four groups enables us to analyze the effect of arousal and valence both separately and interactively.

What emerges from this and other studies with the CEL lists (e.g., Brainerd, Stein, Silveira, Rohenkohl & Reyna, 2008a; Dehon, Larøi & Van der Linden, 2010; Brainerd *et al.*, 2010; Howe, 2007 as only regards the recognition test, so as El Sharkawy, Groth, Vetter, Beraldi & Fast, 2008 and Howe, Candel, Otgaar, Malone & Wimmer, 2010) is that on the one hand, the high arousal increases the memorization of target words (actually presented ones), on the other hand it increases the number of false memories. As regards the valence, it is possible to affirm how the negative valence generally has a greater impact on the memorization process (Brainerd *et al.*, 2008a, 2010; Dehon *et al.*, 2010; El Sharkawy *et al.*, 2008; Howe, 2007; Howe *et al.,* 2010), although a portion of these results could be influenced by the nature of the proposed task (recall or recognition task) (Dehon *et al.*, 2010; El Sharkawy *et al.*, 2008; Howe, 2007; Howe *et al.*, 2010). To return to our example, one thing would be

to present Mariana with a list of words that are related to a concept that is emotionally relevant to her, such as “hospital”. Another thing would be to present her with a list of words all semantically associated with the more neutral concept of "green" (e.g., frog, apple, salad, grass ...): she is supposed to be more likely to be activated by the words related to the hospital sphere and, therefore, to better remember them. However, this activation could be so strong as to lead her to further misremember the lure “hospital” compared to the more neutral lure “green”.

Overview.

On the one hand, nearly all of the research that studied false memory in a bilingual population of children used standard DRM lists, without considering the emotional component, that appears to be lacking also concerning general processing of the emotional lexicon during childhood. The emotional content of the words could be a key to assess semantic processing, since the emotionality of a word is able to activate more the semantic network and consequently a gist-based processing, which would be responsible for false memory.

On the other hand, all of the experiments that intended to study the role of emotional word processing in bilinguals focused on correct memorization and not on memory distortions.

In light of these gaps we found in literature, we would like to investigate whether the connotation of a word in a bilingual second language acts on semantic and mnestic processing as would be for monolingual peers in their native languages. For this purpose, we employed an emotional version of the DRM paradigm to test bilingual and monolingual children.

Aims of the study

The aim of the present study is to examine semantic processing of emotional words (divided according to their arousal and valence) in bilingual children speaking a minority language, as measured in their L2. On the basis of literature, we can postulate that children of the age of our general sample (Mean age = 10.12) would have an emotional semantic processing pattern similar to that of the adults: this means we expect children to produce more false memories for high emotional words. Our explorative scope is to verify if bilingual children show the same pattern of their monolingual peers, with a false memory enhancement for higher emotionally charged lures. We can argue that this result would imply that bilingual semantics would be equally developed as for the monolinguals.

With respect to accuracy rates in the recognition task, we expect similar performances between the two groups, according to literature which certified that bilingualism does not impact on memory performance per se (e.g., Bonifacci, Giombini, Bellocchi & Contento, 2011; Bonifacci, Lombardo, Pedrinazzi, Terracina & Palladino, 2020; Engel de Abreu, 2011; Feng, 2008). To conclude, we would like to consider some control measures that we think could act on connotative semantic performance.

Method

Participants.

A sample of 129 children from two school levels, IV and VI grades, took part in the study (see Table 1.1.). It was recruited at a comprehensive Institute of Pavia, a town in Northern Italy. Teachers report a high degree of bilingualism and a massive presence

of immigrant children in all the classes of the Institute, that covers schooling from kindergarten to first grade secondary school.

Seven children were excluded for learning disabilities (diagnosed by the Public Health Service) or for scoring less than two standard deviations below the expected mean in the vocabulary and reading test: for this purpose, we administered a reading word task and a non-word task (DDE-2, 1995), where children had to read aloud respectively lists of content words and lists of non-words that sound similar to Italian words; fluency and accuracy were taken into account. As regards fluency, it was assessed according to the reading time of each list, measured in syllables per second: the final fluency score consists of the sum of the three word lists scores and, separately, of the three non-words lists scores. As regards accuracy, it consisted of the total amount of errors made by the children (each word read wrongly can contain at most one error; in light of this each word can only be read totally correctly or totally wrongly): errors regard pronunciation of each word and non-word.

Thus, 122 children have been included in the analyses.

Half of the children were bilingual (n = 56), speaking a minority language (30 % of the sample speaking Arabic; 10 % speaking Chinese, 27 % speaking South American Spanish, 7 % speaking Albanian, 5 % speaking Romanian, 3 % speaking Polish, 4 % speaking Brazilian Portuguese and 1 % speaking Turkish and 13 % speaking other dialects) in their family context as L1 and attending Italian language school, so that they were fluent in Italian, their L2. They had not received previous schooling in a language different from Italian. 11 children were born in a country different from Italy (19.6 % of the bilingual sample): the 80.4 % of the sample was made up of second generation immigrant children, attending all their schooling in Italy, from kindergarten to date; as regards the remaining first generation immigrant

children, they all started primary school in Italy. The whole bilingual sample was made up of balanced bilinguals, highly proficient in Italian, in accordance with the average expected for their age, as demonstrated by vocabulary, comprehension and reading scores. Furthermore, we administered two subscales of the Italian version of parental questionnaire ALDeQ (Bonifacci, Mari, Gabbianelli, Ferraguti, Montanari, Burani & Porrelli, 2016): Scale A, "Early milestones" aimed at collecting information regarding the early developmental steps, as first words/sentences, whereas Scale B "Child's current L1 and L2 abilities" aimed at assessing the child's level in each of the two languages spoken. The maximum score that can be obtained in each of the two scale is 18 points. Bilinguals scored 13.51 at Scale A and 8.56 at scale B.

The monolingual group was made up of children with Italian as L1, with comparable age and schooling. (Mean age of the monolingual group: 10.20; Mean age of the bilingual group: 10.04. $p = 0.394$)

Table 1.1. *Descriptive Statistics*

Class	***n***	**Age**			**Genre %**		**%**	
		Range	Mean	SD	Male	Female	Bilinguals	Monolinguals
IV	73	9-10	9.29	1.14	46.58	53.42	50.68	49.32
VI	49	11-12	11.38	0.49	63.27	36.73	38.78	61.22
Tot.	122	9-12	10.12	0.51	53.28	46.72	45.90	54.10

Minority languages are the ones spoken by the minority of the population of a territory and those with lower socioeconomic status; a broad debate related to the relation between bilingual cognition and socioeconomic status (SES. Morton & Harper, 2007; Nair, Biedermann & Nickels, 2017) stressed the necessity to always take into account the role of SES when speaking about bilingualism and considering the associated distinction of minority vs. majority languages.

In light of this, we referred to Hollingshead questionnaire *Four factor Index of Social Status* (1975) in order to calculate the socioeconomic status (SES) of each child, intended as an average of the educational level + profession of the mother and the educational level + profession of the father. We referred to the index and values proposed by the author: it was significantly different between the two groups, as can be seen in Table 1.2. Therefore, SES was treated as a covariate in the statistical analyses.

Table 1.2. *Socioeconomic status between groups*

	Monolinguals (n = 66)	Bilinguals (n = 56)
Low (8-19)	9.8%	30 %
Medium Low (20-29)	18 %	34 %
Medium (30-39)	19,67 %	18 %
Medium High (40-45)	32.78 %	12 %
High (55-56)	19.67 %	6 %
Mean range	40.89 (Medium High)	28.88 (Medium Low)
Standard Deviations	12.51	12.02

** 1 The comparison between SES in the two distributions is statistically significant, t = -4.7501, df = 120.86, p = 5.651e-06**

Materials.

We administered four subtests from Italian standardized batteries, as control measures. (PMA Verbal Meaning, Thurstone, Thurstone & Formaggio, 1957; BVN 5-11 Categorical fluency, Bisiacchi Cendron, Gugliotta, Tressoldi & Vio, 2005; BVN 5-11 Forward Span and Backward Span, Bisiacchi *et al.*, 2005). We chose to aggregate them in two main components that will be of interest for our analyses: vocabulary and working memory.

Vocabulary

In order to assess receptive vocabulary skills, we proposed a verbal meaning task (PMA, 1957) in which participants had to correctly mark the synonym of a given word, choosing between four alternatives. They had 8 minutes to complete the task, composed by 30 trials, as in its original version; scoring consists in counting the number of correct responses.

We administered a fluency task (BVN 5-11, 2005) to assess productive vocabulary: children had to name as many words as possible belonging to a given category (i.e., "fruits": the child had to name as many fruits as possible); they had one minute for each of four categories, as in the test manual. We chose to use this task instead of other similar tasks like Peabody Picture Vocabulary Test (PPVT-III, Dunn & Dunn, 1997) in order to avoid an excess of overlap (collinearity) between our main semantic DRM measure and the vocabulary ones: in fact, PPVT is given as a semantic measure (e.g., Catani, Mesulam, Jakobsen, Malik, Martersteck, Wieneke, Thompson, Thiebaut De Schotten, Dell'Acqua, Weintraub & Rogalski, 2013), while categorical fluency is more often treated as a measure that lies on the edge between productive vocabulary and executive functioning (e.g., Varvara, Varuzza, Sorrentino, Vicari & Menghini, 2014).

Working Memory

Direct and Reverse Span (BVN 5-11, 2005) were administered to assess working memory: in the former task the experimenter read aloud blocks composed by increasing sequences of numbers; at the end of each sequence the participant had to repeat the numbers in the same order. The latter task was similar to the previous one, but at the end of the sequence, the participant had to repeat the numbers in the reverse

order. The tasks ended when participants failed to recall two sequences of the same block. We have chosen an oral working memory task to measure this component as detached from reading and writing skills.

Experimental Task

As regards the main task, participants were exposed to 12 emotional DRM lists, presented in Italian, L2 for the bilinguals and L1 for the monolingual (see Supplementary Materials, Table S1). The materials have been adapted and validatedi from the original CEL lists (Brainerd *et al.*, 2008c), following the same testing procedure used by the Authors on a sample with comparable age. In order to adapt and validate the Cornell/Cortland Emotional Lists (CEL, 2008) to Italian we've conducted a pilot study on 25 monolingual children (Mean age = 9.8), different from the ones that took part in our main study. Arousal and valence were assessed on a dichotomous scale (respectively, high/low and negative/positive), backward associative strength (BAS) on a 7 point Likert scale (from 1 = not associated at all to 7 = very associated). We checked the frequencies of the words included referring to CoLFIS Lexical Database (Bertinetto, Burani, Laudanna, Marconi, Ratti, Rolando & Thornton*,* 2005); in addition we asked children to mark their familiarity with the words, from 0 = not at all to 4 = very well. We only included words marked by the children as known from 2 to 4 and from medium to high frequency words. We've compared our results to the original ones and found a correspondence of 82% as regards the valence and a correspondence of 62 % as regards arousal. We've reassigned the lists to the four Arousal * Valence conditions on the basis of the pilot results.

Twelve lists have been chosen from the adapted-to-Italian 32 Cornell/Cortland Emotional Lists (CEL, Brainerd *et al.*, 2008c). Every list presented in the coding phase

was made up of 10 of the 15 original words, ordered from the one with the stronger BAS, to that with the weaker BAS (Backward Association Strength, which is the degree of association of each word of a list with its corresponding lure, as measured on a 7 point Likert Scale). The 5 words with the lowest BAS value, out of the 15 words of each list, were thus eliminated. The 12 lists were 3 from each of the four Arousal * Valence combination: High Arousal * Negative Valence, Low Arousal * Negative Valence, High Arousal * Positive Valence, Low Arousal * Positive Valence.

The recognition task was related to every Arousal * Valence combination and was made up of 21 words: 9 target words (3 from each presented list), the 3 not presented critical lures corresponding to each presented list, 1 related distractor for each presented list (taken by the 12th position of the 15-words lists) and 6 unrelated distractors, two for each of three unpresented lists, belonging to the same Arousal * Valence combination.

Procedure.

Participants took part in two meetings, the first one was collective and the second one individual, both taking place inside the school, during the scheduled time for educational activities. The individual meetings lasted about 45 minutes per participant. The collective meeting lasted about 20 minutes, involving an entire class at a time and provided for the administration of the vocabulary test: the activity was carried out following the methods and times established by the standardized battery. The individual meeting was in turn divided into two parts: during the first part the children carried out the working memory test, the reading test and the categorical fluency test in paper form, following the methods and times established by the respective standardized batteries. Breaks were foreseen between one task and the next, in order to

avoid fatigue. In the second part, the DRM task was proposed on a Macbook Air laptop computer, with a 13-inch screen and it was run on Open Sesame 2.0 (Mathôt, Schreij & Theeuwes, 2011), in a both visual and auditory modality. Stimuli were written in the font "mono", 46 px, at a resolution of 1024 px X 768 px, with a mean of approximately 70° as visual angle. The audio files were .wav, created with the text to speech program TextAloud 3.0.108 (https://textaloud.it.uptodown.com/windows). The procedure was the same adopted and described by Brainerd *et al.* (2010): each child was instructed to focus on the screen and try to memorize as many presented words as possible. 12 adapted to Italian CEL lists were run. There were four coding sessions, one for each Arousal * Valence combination: after the first coding session there was a 30-sec interval in which children were involved in an easy mathematical task. Following the math distractor task, the first recognition task was presented, in which the child had to press the space bar every time s/he thought s/he recognized the word as part of the previously memorized ones. This procedure was repeated for each of the four sessions, from the coding to the recognition phase: between one block and the next there were breaks of two minutes during which the experimenters asked students short questions to assess their understanding and asked them for feedback regarding perceived difficulty and fatigue.

The order of the two testing phases - the collective and individual one - was randomized between classes, while the order of the two individual parts - paper one and the DRM task - was randomized between participants, in the same way as for the order of the DRM sessions.

Signal detection index.

Signal detection theory is widely used in the literature that refers to DRM paradigm (Rotello, 2017): d' is the index used in order to separate memory performance from noise (which is response bias). Our 'hits' were the correctly recognized words, while our 'false alarms' were the critical lures. The d' indexes obtained (z(H) – z(FA)) were the ones considered in the analyses, so as to work on a clean measure (Fischer & Milfont, 2010), where z(H) was the normalized number of Hits, that in our case were the correctly recognized words (accuracy index) and z(FA) was the normalized number of False Alarms that in our case corresponded to the number of false memories. Here we report the d' measures for the Total Sample (Mean d' = -2.722352e-16, SD = 1.007), for the Bilingual Sample (Mean d' = -0.02623643, SD = 1.0136) and for the Monolingual Sample (Mean d' = 0.021, SD = 1.004): t tests run on these values showed that these means are comparable (Difference between the monolingual d' Mean and the bilingual d' Mean: $p = .614$), as will be better discuss in the following sections.

Results

Statistical analyses are reported below. The core of the research was to examine semantic processing of emotional words (divided according to their valence and arousal) in minority language bilingual children, compared with their monolingual peers. For this purpose, we run on RStudio 1.0.153 (RStudio Team, 2020) a Linear Mixed Model (LMM), which enables the inclusion of random effects, in addition to the fixed effects, in order to account for individual variability in parameters estimate. As can be observed in Table 1.3, our dependent variable was d', while bilingualism, valence and arousal constituted our fixed effects; SES, vocabulary and working memory were used as covariates, in order to take into account a possible role of these variables on d' values.

Table 1.3. *Mixed Model*

	Sum Sq	Mean Sq	Num DF	DenDF	F value	*p* value
Bilingualism	2.0699	2.0699	1	111.19	2.6673	.105
Valence	0.0166	0.0166	1	114.06	0.021	.884
Arousal	19.7976	19.7976	1	114.18	25.5112	.000***
SES	11.4488	11.4488	1	111.02	14.7530	.000 ***
Vocabulary	0.0486	0.0486	1	111.02	0.0626	.803
Working Memory	2.7108	2.7108	1	111.02	3.4931	.064
Bilingualism*Valence	0.3145	0.3145	1	114.06	0.4053	.526
Bilingualism*Arousal	0.1216	0.1216	1	114.18	0.1567	.693
Valence*Arousal	11.7862	11.7862	1	114.19	15.1878	.000 ***
Bilingualism*Valence*Arousal	1.9337	1.9337	1	114.19	2.4917	.117

** 2 Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1*

Subjects and words were our random intercepts (See Table 1.4).

Table 1.4. *Random effects*

Groups Name	Variance	SD	Corr.
Participant (Intercept)	0.08222	0.2867	
Participant (ValencePositive)	0.13644	0.3694	-0.37
Participant (ArousalLow)	0.02010	0.1418	-0.26 (0.13)
Residuals	0.77604	0.8809	

Number of obs: 464, groups: Participants, 116

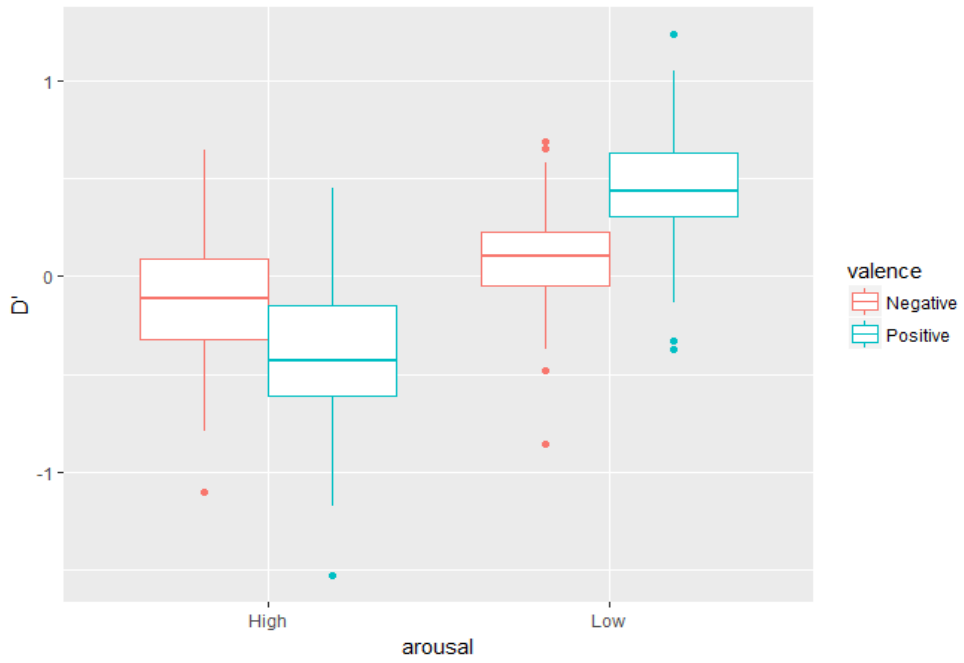
At first, we report the results regarding the emotional semantic processing pattern on the general sample, of both bilinguals and monolinguals. As postulated,

arousal had a statistically significant main effect on d', F (1, 114.18) = 40.547, $p <$ .001: high arousal lures (Mean = - .274) were more prone to be recognized as already seen than the low arousal lures (Mean = 0.263).

Valence did not show a statistically significant main effect, F (1, 113.96) = 0.194, p = .884.

However, the Valence * Arousal interaction was statistically significant, F (1, 114.19) = 13.250, $p <$ .001, as can be observed in Fig. 1.1: false memory production was affected by the Arousal * Valence interaction. To disentangle the interaction effect, we tested differences by using post-hoc multiple comparisons with Tukey HSD corrected p-values. In the low arousal condition, positive valence stimuli (Mean = 0.433) showed significantly higher d' compared to stimuli with negative valence (Mean = 0.093; t(222) = -2.758, p = 0.006), while there was a statistical difference approaching significance in the high arousal condition, in the opposite direction (Mean pos = -0.3899, Mean neg = -0.1575; t(222) = 1.887, p = .061), with positive stimuli showing lower d' compared to negative stimuli.

Figure 1.1. The role of Arousal and Valence on d' values.



As regards our explorative scope, bilingualism was not statistically significant, neither as a main effect, F (1, 111.19) = 2.667, p = .105, nor in the interaction effects: two-way, Bilingualism * Valence, F (1, 114.06) = 0.406, p = .526, Bilingualism * Arousal, F (1, 114.18) = 0.157, p = .693, three-way, Bilingualism * Valence * Arousal, F (1, 114.19) = 2.492, p = .117.

With respect to the accuracy rates, as expected bilingualism proved not to be significant in affecting accuracy scores (p = .549).

With respect to the role played by the control measures, in the Tables below it is possible to observe their mean values and the correlations between variables.

Table 1.5. *Between groups means and standard deviation in the control measures with significance degrees*

	Fluency	*SD*	Direct Span	*SD*	Reverse Span	*SD*	PMA	*SD*
Bilinguals	43,28	9.61	4	1.14	3	0.77	24	6.71
Monolinguals	44,72	8.69	4,75	0.97	3	0.83	26,88	4.77
T and p-value	*t* = -2.1991 *p* =0.030		t = -2.6445 *p* = 0.009*		t = -0.7025 *p* = 0.484		t = -0.7025 *p* =0.483	

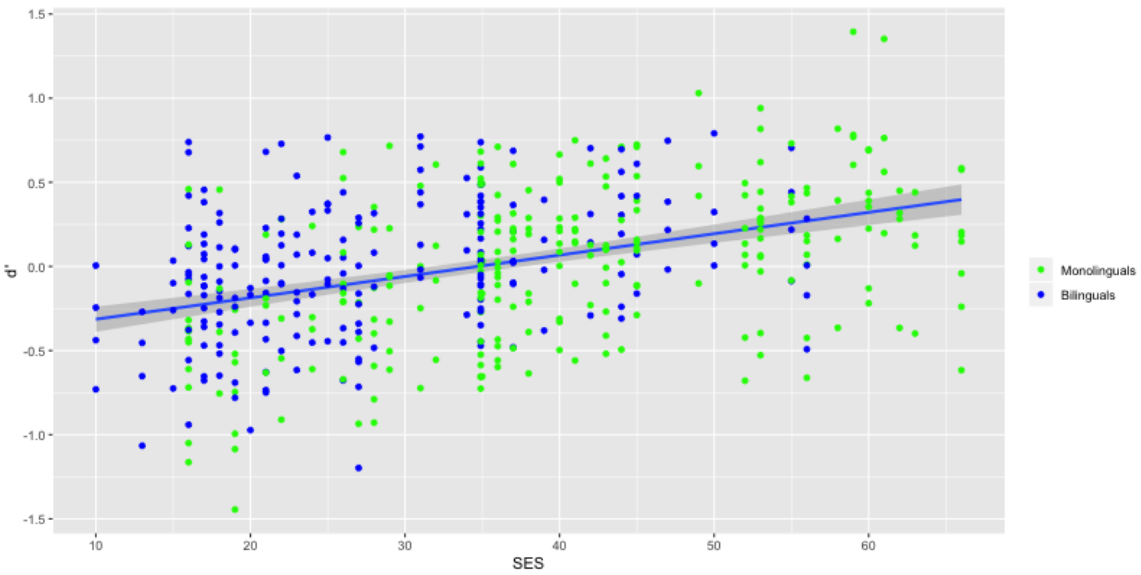
Table 1.6. *Between variables correlation matrix*

** = significative*

	Bilingualism	SES	Vocabulary	Working Memory
Bilingualism	X	-0.44*	-0.32*	-0.19*
SES		X	0.38*	0.02
Vocabulary			X	0.40*
Working Memory				X

More in detail, the socioeconomic status was significant in affecting d' mean values, F (1, 111.02) = 14.753, p < .000. This implies that participants with higher SES had higher accuracy rates in the DRM task.

Figure 1.2. The role of SES on bilinguals and monolinguals



Vocabulary did not show significant effects on d' values, F (1, 111.02) = 0.063, p = .803.

Working memory showed a trend toward statistical significance, F (1, 111.02) = 3.493, p = .064. High scores in the two memory tasks marginally increase the accuracy rates in the DRM task.

Discussion

In this study we focused on how semantics develops in the second language when it is spoken by minority language bilingual children, who learned it mainly within school and the enlarged social context. We aimed in particular to understand how the emotional connotation of the lexicon (its positive or negative valence and its activation – arousal) affects the quality of its memory, by investigating the production of false memories: indeed, false memory proved to be a valid index of semantic activation, as it is the result of a *gist-based* elaboration of the proposed material (Brainerd *et al.*, 2008b).

We meant to explore a domain that appears to be a gap in literature: some authors (e.g., Ayçiçeği & Harris, 2009; see also Eilola, Jelena & Sharma, 2007 for contrasting data) examined the bilingual mnestic processing of emotionally charged verbal materials focusing on true recall or recognition of actually presented words; the authors that examined bilingual false memories didn't consider the role played by the emotional connotation.

Going through our aims, first of all we meant to replicate the results found in literature related to monolinguals, hypothesizing that semantic false memory for higher emotional lures would have been larger than for less emotionally charged lures (e.g., Brainerd *et al.*, 2008c; Dehon *et al.,* 2010; Brainerd *et al.*, 2010; Howe, 2007 as only regards the recognition test, so as El Sharkawy *et al.*, 2008 and Howe *et al.*, 2010). From our results, it emerged that the arousal dimension was significant in affecting the production of false memories: high arousal words were easily recognized and generated a greater number of false memories; this aspect is consistent with the aforementioned literature.

On the other hand, valence was not statistically significant in affecting semantic elaboration, but the Valence * Arousal interaction did: these two emotional dimensions

cannot, undoubtedly, be considered as a single variable that acts indiscriminately on semantics and, through the latter, on processing, but they seem to have a mutual and determinant influence on it.

Our explorative aim was to verify how the bilingual children elaborate emotionally charged words in their L2 compared to their monolingual peers. In particular, we meant to compare the two groups of children, relatively to their pattern of false memories for lures divided according to their arousal and valence (as in Brainerd *et al.*, 2010).

Bilingual children seemed to show the same pattern of results as regards how the emotional component of the elaborated lexicon acts on memory distortion, despite the fact that, due to analysis carried out on the global sample, it is odd to make real inferences and it is only possible to make hypotheses observing a global pattern that goes beyond the "bilingualism" factor.

In particular, arousal proved to be significant in the total group in affecting false memory production: higher arousal lures were more prone to be falsely recognized as already seen. We tested the main effect of arousal also exclusively on the bilingual group and found significance ($p < .001$). This result doesn't seem to support the hypothesis according to which bilingual people might experience their L2 as less emotionally charged (Ayçiçeği & Harris, 2009), and therefore, they might be less prone to let the emotional component of the words distort the elaboration of the lexicon.

Furthermore, there were no differences as regards valence and Valence * Arousal interaction: also considered separately in this group, the former was not significant in affecting false memories, but it was significant in its interaction with the arousal.

These results lead us to consider the semantic ability of the minority language bilingual children as equally developed compared to that of their monolingual peers. In

particular, it emerged that this dimension is developed enough to be affected by the connotational aspect of the presented lexicon, influencing the memory performance and the elaboration of verbal material as much as happens in the monolingual children.

We expected to find similar accuracy rates between monolingual and bilingual children in the correct recognition, according to literature which certified that bilingualism does not act on memory performance (Bonifacci *et al.*, 2011; Engel de Abreu, 2011; Feng, 2008): indeed, what emerged from our results was that there were no statistically significant differences between the two groups as regards the correct recognition of the actually presented words.

These results have been considered in light of our control variables, which proved to be intrinsically related with both the concept of semantics and of bilingualism.

In particular, it emerged that vocabulary skills did not affect the memorization and correct performance in the recognition task. The correlation found between vocabulary and bilingualism was coherent with the literature (Bialystok, 2009), confirming that the bilingual children in this age group (Mean age of the monolingual group: 10.20; Mean age of the bilingual group: 10.04) seem to have a less developed vocabulary knowledge if compared to that of their monolingual peers and if measured in isolation in each of their two spoken languages. These results lead us to consider semantic development as detached from the vocabulary component, going in the direction of the literature that considers the bilingual semantic capacity as language-independent and as a shared system between the two languages spoken (e.g., Dong *et al.*, 2005; Francis, 2005; Navracsics, 2002; van Heuven *et al.*, 1998).

Working memory proved to have a weight in the production of false memories: better results in the working memory tasks seem to be associated to better accuracy rates in the DRM recognition phases. Although considering the combined index it is only a

trend towards significance, this result appears interesting since it provides evidences regarding the different mechanisms underlying the production of false memories: even if the role of semantic processing as the primary architect of the production of this type of false memories is established, a greater understanding of how much is due to semantics and how much, instead, is due to working memory can be useful to make the DRM task increasingly more accurate: a good support in working memory is required to retain the already repeated items and to pre-represent the ones that have to be further produced. As Baddeley, Allen and Hitch (2011) postulated in their last refined working memory model, working memory interacts with long term memory: this interaction can explain much about the final working memory performance. In light of these considerations, we think that the relation between working memory and DRM recognition deserves to be further explored.

The socioeconomic status proved to be significant in affecting false memory: higher levels of SES would correspond to an higher number of false memories and, therefore, to higher semantic levels. This result goes in the direction of the literature that highlighted SES as a predictor of lexical, semantic and vocabulary knowledges (e.g., Gatt, Baldacchino & Dodd, 2020; Maguire, Schneider, Middleton, Ralph, Lopez, Ackerman & Alyson, 2018), in the relation between linguistic abilities and school performance (e.g., von Stumm, Rimfeld, Dale & Plomin, 2020) and in explaining working memory abilities (e.g., Tine, 2014). It is important to notice that the SES variable in our study was highly and negatively correlated with bilingualism (- .39, $p < .001$), which means that the bilingual children included in our sample had a lower socioeconomic status compared to their monolingual peers. These results lead us to propose that although minority language bilingualism is often associated with a condition of lower socio-economic status, these two variables can be considered in light of the

different influence they seem to have on semantics: if coming from a family with high socioeconomic status seems to positively influence the semantic competence, the same cannot be said of the condition of bilingualism.

The main limit attributable to the study concerns the inability to emphasize the specificity of all the linguistic and cultural backgrounds of the bilingual sample, made up of the broad definition of minority language bilingualism, but which actually involved children from a large number of countries and nationalities, very different one from each other and, furthermore, which involved extremely heterogeneous linguistic realities. The mediation carried out by teachers and experimenters did not solve the impossibility of

adapting the materials to all the linguistic realities present within the involved sample.

Moreover, despite this aspect could go beyond the scope of this research, while interpreting the results of the study, one cannot avoid considering the role played by the belonging culture and the homogeneity of the here considered bilingual sample as a mediator of the emotional attribution. Although these children were prematurely exposed to their L2 at an immersive level, most of them had experiences that differed from the ones expressed by the Italian culture (the one of their L2); this aspect can have a weight in the subjective emotional interpretation that the children gave to the proposed words. One further thing that can be mentioned relatively to the influence of cultural differences on memory distortion is schemas, defined by Webb, Turney and Dennis (2016) as "useful memory tools that allow information to be categorized according to a common concept or theme", adaptable associative networks of knowledge extracted over multiple similar experiences (Robin & Moscovitch, 2017). Schemas consist of a collection of "underlying personal beliefs, social pressures, biases and heuristics, and cultural assumptions" which, taken together, permeates our mental organization and constitute the basis for our daily

experience and a key access to interpret the world (Hendricks, 2021). In light of this implicit process, it is clear that not all schemas are universally accepted and that they strongly depend on the environment in which one is immersed. These differences play a role on memory distortions and the priority, the relevance that one put in his elaboration largely depend on them (Webb *et al.*, 2016).

To date, we are working on extending the project to different bilingual samples (i.e., of majority language), as well as different age groups (seniors aged 65 and over). It would also be interesting to consider the reaction times associated with the recognition task, evaluating how they change on the basis of arousal and valence. Another possibility would be to implement the study with the joint use of techniques, such as eye tracking and galvanic skin response (GSR), the former aimed at detecting the direction of the gaze, the amount of saccadic responses and the duration of the fixations while the DRM lists are presented and recognized, the latter to detect physiological activation in relation to the emotional nature of the presented stimuli.

Conclusions

From our results it appears that children from 9 to 11 years produce false memories, as an expression of semantic processing, similarly to adults. Equally, emotion plays a similar role on the aforementioned semantic processing, with high arousing lures being more false recognized than low arousal ones. Minority language bilingual children seem to develop the connotative dimension of semantics in their L2 equally to their monolingual peers. This similarity has been shown also as regards accuracy rates in the recognition task, as a proof that bilingualism doesn't affect memory performance per se. Semantic processing seems to be detached from vocabulary knowledge, usually negatively affected by bilingualism in this age period (e.g., Bialystok, 2009). To conclude, an important role on false memory's production seems to be played by socioeconomic status, negatively

correlated with bilingualism: this aspect strengthens the need to consider bilingualism in its entirety, linked to its social context and to its various features.

The debate about bilingual semantics is still open and it still fascinates many researchers. Semantics is a sort of bridge between two of the dimensions identified by Bialystok (2009) as crucially sensitive for bilingualism understanding: vocabulary and memory. A greater understanding of bilingual semantics in relation to these two dimensions can help us to disambiguate the place it has in relation to language, to better clarify how dependent it is on the development of the lexicon and how much it affects its memorization.

Although the presented study cannot and does not aim at exhausting a debate as broad, complex and multifaceted, nevertheless it opens the way for an unprecedented look at these fundamental themes. Here, we aim to consider them not within an abstract void, but as inseparably imbued with emotional colours, which permeate, shape and determine the movements and mutual relationships.

References.

Altarriba, J., & Basnight-Brown, D. M. (2009). An overview of semantic processing in bilinguals: Methods and findings. In A. Pavlenko (ed.), *The bilingual mental lexicon: Interdisciplinary approaches* (pp. 79–99). Clevedon: Multilingual Matters. DOI: 10.22158/eltls.v2n3p43

Anastasi, J., Rhodes, M., Marquez, S., & Velino, V. (2005). The incidence of false memories in native and non-native speakers. *Memory*, *13*, 815–828. DOI: 10.1080/09658210444000421

Anooshian, L. J., & Hertel, P. T. (1994). Emotionality in free recall: Language specificity in bilingual memory. *Cognition & Emotion*, *8*, 503–514. DOI: 10.1080/02699939408408956

Ayçiçeği, A., & Harris, C. L. (2004). Brief Report: Bilinguals' recall and recognition of emotion words. Cognition and Emotion, 18, 977–987. DOI: 10.1080/02699930341000301

Ayçiçeği, A., & Harris, C. L. (2009). Emotion-memory effects in bilingual speakers: A levels-of-processing approach. *Bilingualism: Language and Cognition*, *12*, 291–303. DOI: 10.1017/S1366728909990125

Baddeley, A. D., Allen, R. J., & Hitch, G. J. (2011). Binding in visual working memory: the role of the episodic buffer. *Neuropsychologia*, 49, 1393–1400. DOI: 10.1016/j.neuropsychologia.2010.12.042

Baugerud, G. A., Howe, M. L., Magnussen, S., & Melinder, A. (2016). Maltreated and non-maltreated children's true and false memories of neutral and emotional word lists in the Deese/Roediger–McDermott task. *Journal of Experimental Child Psychology*, *143*, 102–110. DOI: 10.1016/j.jecp.2015.10.007

Bertinetto, P. M., Burani, C., Laudanna, A., Marconi, L., Ratti, D., Rolando, C., & Thornton A. M. (2005). Corpus e Lessico di Frequenza dell'Italiano Scritto (CoLFIS*)*.

Bialystok, E. (2009). Bilingualism: The good, the bad, and the indifferent. *Bilingualism: Language and Cognition, 12,* 3–11. DOI: 10.1017/S1366728908003477

Binet, A. (1900). La suggestibilité (Vol. 3). Librairie C. Reinwald.

Bisiacchi, P., Cendron, M., Gugliotta, M., Tressoldi, P. E., & Vio, C. (2005). BVN 5–11. *Batteria Di Valutazione Neuropsicologica per l'Età Evolutiva*. Trento: Erickson.

Bjorklund, D. F. (2000). *False-memory Creation in Children and Adults: Theory, Research, and Implications* (1st ed.). Psychology Press.

Bonifacci, P., Giombini, L., Bellocchi, S., & Contento, S. (2011). Speed of processing, anticipation, inhibition and working memory in bilinguals. *Developmental Science, 14*, 256–269. DOI: 10.1111/j.1467-7687.2010.00974.x

Bonifacci, P., Lombardo, G., Pedrinazzi, J., Terracina, F., & Palladino, P. (2020). Literacy Skills in Bilinguals and Monolinguals with Different SES. *Reading and Writing Quarterly, 36*, 243-259. DOI: 10.1080/10573569.2019.1635057

Bonifacci, P., Mari, R., Gabbianelli, L., Ferraguti, E., Montanari, F., Burani, F., & Porrelli, M. (2016). Sequential bilingualism and specific language impairment: The Italian version of ALDeQ parental questionnaire. *BPA-Applied Psychology Bulletin* (Bollettino di Psicologia Applicata), *64*, 50–64.

Brainerd, C. J., Holliday, R. E., Reyna, V. F., Yang, Y., & Toglia, M. P. (2010). Developmental reversals in false memory: Effects of emotional valence and arousal. *Journal of Experimental Child Psychology, 107*, 137–154. DOI: 10.1016/j.jecp.2010.04.013

Brainerd, C. J., Reyna, V. F., & Forrest, T. J. (2002). Are Young Children Susceptible to the False-Memory Illusion? *Child Development*, *73*, 1363–1377. DOI: 10.1111/1467-8624.00477

Brainerd, C. J., Reyna, V. F., & Holliday, R. E. (2018). Developmental reversals in false memory: Development is complementary, not compensatory. *Developmental Psychology*, *54*, 1773–1784. DOI: 10.1037/dev0000554

Brainerd, C. J., Stein, L. M., Silveira, R. A., Rohenkohl, G., & Reyna, V. F. (2008a). How Does Negative Emotion Cause False Memories? *Psychological Science*, *19*, 919–925. DOI: 10.1111/j.1467-9280.2008.02177.x

Brainerd, C. J., Yang, Y., Reyna, V. F., Howe, M. L., & Mills, B. A. (2008b). Semantic processing in "associative" false memory. *Psychonomic Bulletin & Review*, *15*, 1035–1053. DOI: 10.3758/PBR.15.6.1035

Brainerd, C. J., Yang, Y., Toglia, M. P., Reyna, V. F., & Stahl, C. (2008c). Emotion and false memory: The Cornell/Cortland norms. In *annual meeting of the Psychonomic Society, Chicago, IL.*

Brase, J., & Mani, N. (2017). Effects of learning context on the acquisition and processing of emotional words in bilinguals. *Emotion*, *17*, 628–639. DOI: 10.1037/emo0000263

Braunstein, V., Ischebeck, A., Brunner, C., Grabner, R. H., Stamenov, M., & Neuper, C. (2012). Investigating the influence of proficiency on semantic processing in bilinguals: An ERP and ERD/S analysis. *Acta neurobiologiae experimentalis, 72*, 421-38.

Bruck, M., & Ceci, S. J. (1999). The suggestibility of children's memory. *Annual Review of Psychology*, *50*, 419–439. DOI: 10.1146/annurev.psych.50.1.419

Catani, M., Mesulam, M. M., Jakobsen, E., Malik, F., Martersteck, A., Wieneke, C., Thompson, C. K., Thiebaut De Schotten, M., Dell'Acqua, F., Weintraub, S., & Rogalski, E. (2013). A novel frontal pathway underlies verbal fluency in primary progressive aphasia. *Brain*, *136*, 2619–2628. DOI: 10.1093/brain/awt163

Chen, P., Lin, J., Chen, B., Lu, C., & Guo, T. (2015). Processing emotional words in two languages with one brain: ERP and fMRI evidence from Chinese–English bilinguals. *Cortex*, *71*, 34–48. DOI: 10.1016/j.cortex.2015.06.002

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology, 58*, 17–22. DOI: 10.1037/h0046671

Dehon, H., Larøi, F., & Van der Linden, M. (2010). Affective valence influences participant's susceptibility to false memories and illusory recollection. *Emotion, 10*, 627–639. DOI: 10.1037/a0019595

Dong, Y., Gui, S., & Macwhinney, B. (2005). Shared and separate meanings in the bilingual mental lexicon. *Bilingualism: Language and Cognition*, *8*, 221–238. DOI: 10.1017/s1366728905002270

Dunn, L. M., & Dunn, L. M. (1997). *PPVT-III: Peabody picture vocabulary test.* Circle Pines, MN:American Guidance Service

Eilola, T., Jelena, H., & Sharma, D. (2007). Emotional activation in the first and second language. *Cognition and Emotion, 21*, 1064–1076. DOI: 10.1080/02699930601054109

El Sharkawy, J., Groth, K., Vetter, C., Beraldi, A., & Fast, K. (2008). False memories of emotional and neutral words. Behavioural Neurology, 19, 7–11. DOI: 10.1155/2008/587239

Engel de Abreu, P. M. J. (2011). Working memory in multilingual children: Is there a bilingual effect? *Memory*. *19*, 529–537. DOI: 10.1080/09658211.2011.590504

Feng, X. (2008). *Working memory and bilingualism: An investigation of executive control and processing speed*. Toronto, ON, Canada: York University.

Ferré, P., García, T., Fraga, I., Sánchez-Casas, R., & Molero, M. (2010). Memory for emotional words in bilinguals: Do words have the same emotional intensity in the first and in the second language? *Cognition & Emotion*, *24*, 760–785. DOI: 10.1080/02699930902985779

Fischer, R., & Milfont, L. T. (2010). Standardization in psychological research. *International Journal of Psychological Research*, *3*, 88–96. DOI: 10.21500/20112084.852

Floccia, C., Delle Luche, C., Lepadatu, I., Chow, J., Ratnage, P., & Plunkett, K. (2020). Translation equivalent and cross-language semantic priming in bilingual toddlers, *Journal of Memory and Language*, *112,* 104086. DOI: 10.1016/j.jml.2019.104086

Francis, W. S. (2005). Bilingual Semantic and Conceptual Representation. In J. F. Kroll & A. M. B. De Groot (eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 251–267). Oxford University Press.

Francis, W. S. (2018). Shared core meanings and shared associations in bilingual semantic memory: Evidence from research on implicit memory. *International Journal of Bilingualism*, *24*, 464–477. DOI: 10.1177/1367006918814375

Gathercole, V. C. M. (2020). Preface: Special issue on semantic representation and processing in bilinguals. *International Journal of Bilingualism*, *24*, 457–463. DOI: 10.1177/1367006918814374

Gatt, D., Baldacchino, R., & Dodd, B. (2020). Which measure of socioeconomic status best predicts bilingual lexical abilities and how? A focus on four-year-olds exposed

to two majority languages. *Journal of Child Language*, *47*, 737–765. DOI: 10.1017/S0305000919000886

Grabovac, B., & Pléh, C. (2014). Processing of the emotional value of words in an emotional stroop task by late Hungarian-Serbian bilinguals. *Magyar Pszichológiai Szemle*, *69*, 731–745. DOI: 10.1556/mpszle.69.2014.4.5

Graves, D. F., & Altarriba, J. (2014). False Memories in Bilingual Speakers. *Foundations of Bilingual Memory*, 205–221. DOI: 10.1007/978-1-4614-9218-4_10

Hendricks, Z. (2021). Cross-Cultural Differences in Memory, Beliefs, and Mental Schemas. *Inquiries Journal*, 13.

Hollingshead, A. B. (1975). Four-factor index of social status. Unpublished manuscript, Yale University, New Haven, CT.

Howe, M. L. (2007). Children's Emotional False Memories. *Psychological Science*, *18*, 856–860. DOI: 10.1111/j.1467-9280.2007.01991.x

Howe, M. L., Candel, I., Otgaar, H., Malone, C., & Wimmer, M. C. (2010). Valence and the development of immediate and long-term false memory illusions. *Memory*, *18*, 58–75. DOI: 10.1080/09658210903476514

Howe, M. L., Gagnon, N., & Thouas, L. (2008). Development of false memories in bilingual children and adults☆. *Journal of Memory and Language*, *58*, 669–681. DOI: 10.1016/j.jml.2007.09.001

Howe, M. L., Wimmer, M. C., Gagnon, N., & Plumpton, S. (2009). An associative-activation theory of children's and adults' memory illusions. *Journal of Memory and Language*, *60*, 229-251. DOI: 10.1016/j.jml.2008.10.002

Kajcsa, B. (2010). The Organization of Early Bilingual Mental Lexicon in Light of a Picture-Naming Task. *Acta Universitatis Sapientiae, Philologica*, 2, 326-336.

Kaplan, R. L., Van Damme, I., Levine, L. J., & Loftus, E. F. (2015). Emotion and False Memory. *Emotion Review*, *8*, 8–13. DOI: 10.1177/1754073915601228

Kastenbaum, J. G., Bedore, L. M., Peña, E. D., Sheng, L., Mavis, I., Sebastian-Vaytadden, R., Rangamani, G., Vallila-Rohter, S., & Kiran, S. (2019). The influence of proficiency and language combination on bilingual lexical access. Bilingualism: Language and Cognition, 22, 300–330. DOI: 10.1017/s1366728918000366

Khateb, A., Pegna, A. J., Michel, C. M., Mouthon, M., & Annoni, J. M. (2016). Semantic relatedness and first-second language effects in the bilingual brain: A brain mapping study. *Bilingualism: Language and Cognition, 19*, 311–330. DOI: 10.1017/S1366728915000140

Klein, S. B., Robertson, T. E., & Delton, A. W. (2009). Facing the future: Memory as an evolved system for planning future acts. *Memory & Cognition*, *38*, 13–22. DOI: https://doi.org/10.3758/mc.38.1.13

Lecce, S., Zocchi, S., Pagnin, A., Palladino, P., & Taumoepeau, M. (2010). Reading minds: The Relation Between Children's Mental State Knowledge and Their Metaknowledge About Reading. *Child Development, 81*, 1876–1893. DOI: 10.1111/j.1467-8624.2010.01516.x

Lee, T. H., Greening, S. G., & Mather, M. (2015). Encoding of goal-relevant stimuli is strengthened by emotional arousal in memory. *Frontiers in Psychology*, *6*, 1173. DOI: 10.3389/fpsyg.2015.01173

Maguire, M., Schneider, J. M., Middleton, A., Ralph, Y., Lopez, M., Ackerman, R. A., & Alyson, A. (2018). Vocabulary knowledge mediates the link between

socioeconomic status and word learning in grade school. *Journal of Experimental Child Psychology*. *166*, 679–695. DOI: 10.1016/j.jecp.2017.10.003

Marmolejo, G., Diliberto-Macaluso, K., & Altarriba, J. (2009). False memory in bilinguals: Does switching languages increase false memories? *The American Journal of Psychology, 122*, 1–16.

Mathôt, S., Schreij, D., & Theeuwes, J. (2011). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods, 44*, 314–324. DOI: 10.3758/s13428-011-0168-7

Metzger, R. L., Warren, A. R., Shelton, J. T., Price, J., Reed, A. W., & Williams, D. (2008). Do children "DRM" like adults? False memory production in children. *Developmental Psychology*, *44*, 169–181. DOI: 10.1037/0012-1649.44.1.169

Montrul, S., & Foote, R. (2012). Age of acquisition interactions in bilingual lexical access: A study of the weaker language of L2 learners and heritage speakers. *International Journal of Bilingualism*, *18*, 274–303. DOI: 10.1177/1367006912443431

Morton, J. B., & Harper, S. N. (2007). What did Simon say? Revisiting the bilingual advantage. *Developmental Science*, *10*, 719–726. DOI: 10.1111/j.1467-7687.2007.00623.x

Nair, V., Biedermann, B., & Nickels, L. (2017). Effect of socio-economic status on cognitive control in non-literate bilingual speakers. *Bilingualism: Language and Cognition*, *20*, 999–1009. DOI: 10.1017/S1366728916000778

Navracsics, J. (2002). Bilingual semantic representation and lexical access. *Acta Linguistica Hungarica*, *49*, 225–247. DOI: 10.1556/ALing.49.2002.2.3

Newman, E. J., & Lindsay, D. S. (2009). False memories: What the hell are they for?. *Applied Cognitive Psychology, 23*, 1105-1121. DOI: 10.1002/acp.1613

Novitskiy, N., Myachykov, A., & Shtyrov, Y. (2018). Crosslinguistic interplay between semantics and phonology in late bilinguals: Neurophysiological evidence. *Bilingualism: Language and Cognition, 22*, 209–227. DOI: 10.1017/S1366728918000627

Oliveira, H. M., Albuquerque, P. B., & Saraiva, M. (2018). O Estudo das falsas memórias: reflexão histórica. Temas Em Psicologia, 26, 1763–1773. DOI: 10.9788/tp2018.4-03pt

Otgaar, H., Howe, M. L., Muris, P., & Merckelbach, H. (2018). Associative Activation as a Mechanism Underlying False Memory Formation. *Clinical Psychological Science*, *7*, 191–195. DOI: 10.1177/2167702618807189

Otgaar, H., Howe, M. L., Muris, P., & Merckelbach, H. (2019). Dealing With False Memories in Children and Adults: Recommendations for the Legal Arena. *Policy Insights from the Behavioral and Brain Sciences*, *6*, 87–93. DOI: 10.1177/2372732218818584

Otgaar, H., & Smeets, T. (2010). Adaptive memory: Survival processing increases both true and false memory in adults and children. Journal of Experimental Psychology: Learning, Memory, and Cognition, 36, 1010–1016. DOI: 10.1037/a0019402

Puig-Mayenco, E., Cunnings, I., Bayram, F., Miller, D., Tubau, S., & Rothman, J. (2018). Language Dominance Affects Bilingual Performance and Processing Outcomes in Adulthood. *Frontiers in Psychology*, *9*. DOI: 10.3389/fpsyg.2018.01199

Quas, J. A., Qin, J., Schaaf, J. M., & Goodman, G. S. (1997). Individual differences in children's and adults' suggestibility and false event memory. *Learning and Individual Differences*, 9, 359-390. DOI: 10.1016/S1041-6080(97)90014-5

Quas, J. A., Rush, E. B., Yim, I. S., Edelstein, R. S., Otgaar, H., & Smeets, T. (2015). Stress and emotional valence effects on children's versus adolescents' true and false memory. *Memory*, *24*, 696–707. DOI: 10.1080/09658211.2015.1045909

Reyna, V. F., & Brainerd, C. J. (1995). Fuzzy-trace theory: An interim synthesis. *Learning and Individual Differences, 7*, 1–75. DOI: 10.1016/1041-6080(95)90031-4

Reyna, V. F., Corbin, J. C., Weldon, R. B., & Brainerd, C. J. (2016). How fuzzy-trace theory predicts true and false memories for words, sentences, and narratives. *Journal of applied research in memory and cognition, 5*, 1-9.

Reyna, V. F., & Lloyd, F. (1997). Theories of false memory in children and adults. *Learning and Individual Differences*, *9*(2), 95–123. DOI: 10.1016/s1041-6080(97)90002-9

Reyna, V. F., Mills, B., Estrada, S., & Brainerd, C. J. (2007). False memory in children: Data, theory, and legal implications. In M. P. Toglia, J. D. Read, D. F. Ross, & R. C. L. Lindsay (Eds.), *The handbook of eyewitness psychology, Vol. 1. Memory for events* (pp. 479–507). Lawrence Erlbaum Associates Publishers. DOI: 2006-22582-017

Riemer, N. (2010). *Introducing Semantics (Cambridge Introductions to Language and Linguistics)*. UK: Cambridge University Press.

Riesthuis, P., Otgaar, H., & Wang, J. (2019). Differences between Bilinguals and Monolinguals in False Memory Production? A Look into the DRM Paradigm Using Contextual Details. *Psychological Research on Urban Society*, *2*, 88. DOI: 10.7454/proust.v2i2.50

Robin, J., & Moscovitch, M. (2017). Details, gist and schema: hippocampal–neocortical interactions underlying recent and remote episodic and spatial memory. *Current Opinion in Behavioral Sciences*, *17*, 114–123. DOI: 10.1016/j.cobeha.2017.07.016

Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in list. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 803–814. DOI: 10.1037/0278-7393.21.4.803

Rotello, C. M. (2017). Signal detection theories of recognition memory. In J. T. Wixted (ed.), *Learning and Memory: A Comprehensive Reference*, (2nd ed. Vol. 4, pp. 201–225). Amsterdam: Elsevier. DOI: 10.1016/b978-0-12-809324-5.21044-4

RStudio Team (2020). *RStudio: Integrated Development for R. RStudio, PBC, Boston, MA URL http://www.rstudio.com/*

Sanford, A. J. (2006). Semantics in psychology. In K. Brown (ed.), *Encyclopedia of language and linguistics* (2nd ed. Vol. 11, pp. 152–158). Amsterdam: Elsevier. DOI: 10.1016/b0-08-044854-2/01106-8

Schacter, D. L. (2012). Adaptive constructive processes and the future of memory. *American Psychologist, 67*, 603–613. DOI: https://doi.org/10.1037/a0029869

Schacter, D. L., Guerin, S. A., & St. Jacques, P. L. (2011). Memory distortion: an adaptive perspective. *Trends in Cognitive Sciences*, *15*, 467–474. DOI: https://doi.org/10.1016/j.tics.2011.08.004

Schmidtke, J. (2014). Second language experience modulates word retrieval effort in bilinguals: evidence from pupillometry. *Frontiers in Psychology*, *5*. DOI: 10.3389/fpsyg.2014.00137

Seamon, J. G., Luo, C. R., Schlegel, S. E., Greene, S. E., & Goldenberg, A. B. (2000). False Memory for Categorized Pictures and Words: The Category Associates Procedure for Studying Memory Errors in Children and Adults. *Journal of Memory and Language*, *42*, DOI: 10.1006/jmla.1999.2676

Small, M. H. (1896). The suggestibility of children. *The Pedagogical Seminary*, 4, 176-220. DOI: 10.1080/08919402.1896.10534805

Stern, W. (1910). Abstracts of lectures on the psychology of testimony and on the study of individuality. *The American Journal of Psychology*, 21, 270-282. DOI: 10.2307/1413003

Storms, G., Ameel, E., & Malt, B. C. (2015). Development of cross-language lexical influence. *International Journal of Bilingual Education and Bilingualism*, *18*, 529–547. DOI: 10.1080/13670050.2015.1027142

Thurstone, L. L., Thurstone, T. G., & Formaggio, T. (1957). *Batteria fattoriale delle attitudini mentali primarie: 4-6*: *manuale di istruzioni.* Firenze: Organizzazioni Speciali.

Tine, M. (2014). Working Memory Differences Between Children Living in Rural and Urban Poverty, *Journal of Cognition and Development, 15*, 599–613. DOI: 10.1080/15248372.2013.797906

van Heuven, W. J. B., Dijkstra, T., & Grainger, J. (1998). Orthographic neighborhood effects in bilingual word recognition. *Journal of Memory and Language, 39*, 458–483. DOI: 10.1006/jmla.1998.2584

Varvara, P., Varuzza, C., Sorrentino, A. C. P., Vicari, S., & Menghini, D. (2014). Executive functions in developmental dyslexia. *Frontiers in Human Neuroscience*, *8*. DOI: 10.3389/fnhum.2014.00120

von Stumm, S., Rimfeld, K., Dale, P. S., & Plomin, R. (2020). Preschool Verbal and Nonverbal Ability Mediate the Association Between Socioeconomic Status and School Performance. *Child Development, 91*, 705–714. DOI: 10.1111/cdev.13364

Webb, C. E., Turney, I. C., & Dennis, N. A. (2016). What's the gist? The influence of schemas on the neural correlates underlying true and false memories. *Neuropsychologia*, *93*, 61–75. DOI: 10.1016/j.neuropsychologia.2016.09.023

Zhang, W., Gross, J., & Hayne, H. (2016). The effect of mood on false memory for emotional DRM word lists. *Cognition and Emotion*, *31*, 526–537. DOI: 10.1080/02699931.2016.1138930

Zhou, J. (2019). Explaining Theories of Verbal False Memory. *Proceedings of the 3rd International Conference on Culture, Education and Economic Development of Modern Society (ICCESE 2019* DOI: 10.2991/iccese-19.2019.28

In the previous chapter we tested how emotionality conveyed by the words affected verbal false memories' production in bilingual children aged from 9 to 11, in comparison with their monolingual peers. In particular, our bilingual sample was composed by children speaking a minority language as L1 in their family context and Italian as L2, having learned it in through their school and social context. Our exploratory scope was to verify if a second language acquired can reach so much depth as to be influenced by connotation (intending with this term the emotionality conveyed by the meaning of the words) as it has been found in monolingual children. We found that the pattern played by emotionality is comparable between monolingual and bilingual children. This result seems to demonstrate that despite a L2 vocabulary skill that can still be improved, their semantic competence already reached the one of their monolingual peers, maybe benefiting from the competence they have in their L1. This result seems to support the view of an integrated semantic system between spoken languages. Another interesting aspect that emerged from our results was the role played by socioeconomic status. Indeed, it significantly affected verbal false memory performance above being bilingual or not. In particular, children who came from lower SES families were less prone to be influenced by emotionality and, consequently, less prone to acquire that depth of meaning we mentioned above.

This result lead us to the aim of further testing this variable as determinant in affecting literacy at this school level. We are going to expose our study and our results concerning this aspect in the next chapter.

CHAPTER 3

To be or not to be bilingual: How reading comprehension and vocabulary size in children are affected by SES

Under Review: Cangelosi, M., Barichello, C., Dijsktra, T., & Palladino, P. (2022). *Reading and Writing Quarterly Journal*.

Abstract.

Theories of reading in bilingual children have so far neglected to combine the contribution of SES, being bilingual or not, and vocabulary size. We investigated the role of socioeconomic status (SES) in the unfolding of vocabulary and text-comprehension skills in bilingual language minority children, who are educated exclusively in their L2. For comparison, a monolingual control group was also tested, consisting of children of the same age and schooling. We administered a synonyms task to measure vocabulary and an inference test to measure reading comprehension. Socioeconomic status scores were obtained through the Hollingshead Four Factor Index questionnaire. Our results showed that being bilingual did not affect vocabulary and text-comprehension skills in the age group considered (9-11 years). Instead, SES played a significant role in relation to age and schooling. Interestingly, two different mediation analyses demonstrated that vocabulary size accounted for a significant portion of the causal relationship between SES and reading comprehension, and that reading comprehension itself can explain part of the relationship between SES and vocabulary. The social and educational implications of the findings are discussed.

Keywords: SES; Bilingualism; Vocabulary; Comprehension; Children

Introduction

Learning to read is an effortful and long-term process. Many researchers agree that an important goal of learning to read is to acquire efficient reading comprehension (Oakhill, 2020). Reading comprehension is a complex and articulated skill, aimed at combining textual meanings and reader knowledge in a coherent and well-integrated mental model (see, e.g., Kintsch, 1998).

Reading is a regular and important activity for many students in Europe and all around the world who have a family history of immigration. The 2019 Global Education Monitoring Report estimated that in 2017, the number of international migrants in the world amounted to 258 million, around 3.4% of world's population. About 64% if them were in the richest countries where the percentage of migrants compared to the global population passed from 10% in 2000 to 14% in 2017 (UNESCO, 2019). Some of the immigrants, who attend school, learn to read in a second language different from the language spoken at home. Most of them live in a poor socioeconomic environment that may significantly affect their vocabulary size and literacy acquisition (for a review, see August & Shanahan, 2006). Their children may be defined as bilinguals according to a broad definition of being bilingual that includes children who know and use two or more languages in their everyday life (Grosjean, 1992). Though a large body of research has investigated language development and reading acquisition in language minority bilingual children (i.e., children with an L1 spoken by a minority group, August & Shanahan, 2006), less research considers the complex process of reading comprehension in relation to socioeconomic status (SES) and bilingual language skills.

According to the Simple View of Reading (Gough & Tunmer, 1986; Hoover & Gough, 1990), when reading decoding is well acquired (such as from third/fourth school grade), reading comprehension skills overlap rather closely with oral language

comprehension skills. Therefore, it is crucial to examine reading comprehension in late primary school, because this is the period in which decoding is expected to be fully acquired, and the reading comprehension process is less strongly affected by it.

In the present study, we aimed to examine reading comprehension in children in late primary school (fourth to sixth grades) to clarify the roles of SES and being bilingual in affecting vocabulary knowledge. In fact, to clarify their interrelation, we studied how all three variables contribute to reading comprehension performance. We will now examine these variables, starting with vocabulary size and its relation to reading comprehension, and then moving on to SES and being bilingual.

Vocabulary size and reading comprehension

Reading is a complex ability composed of several skills requiring the integration of different types of units (Perfetti & Adlof, 2012). To begin with, readers need to identify individual words and decode them. Next, they must connect single words to form a continuously updated representation of the text (word-to-text integration; Perfetti & Adlof, 2012). Finally, a higher level of reading comprehension is required, where the meaning of each sentence is incorporated throughout reading the text. Here, inference-making and text structure building abilities are needed (e.g., Gernsbacher, 1997).

That vocabulary knowledge plays a role in reading comprehension is a robust finding, replicated by many studies (e.g., Beck, McKeown & Kukan, 2002; Droop & Verhoeven, 2003; Verhoeven & van Leeuwe, 2008; Verhoeven et al., 2011). In most of the studies, vocabulary size was defined as the number of known words (i.e., breadth of knowledge; but see also Perfetti and Hart, 2001). Word knowledge has a direct effect on the ability to understand a text, because word meaning is the basic building block of the larger structure constructed by text-comprehension processes (e.g., Genrnsbacher, 1997; Gensbacher, Robertson, Palladino & Werner, 2020). In fact, when students have limited

vocabulary skills, difficulties arise in both recall and reading comprehension of a text (Beck, Perfetti, & McKeown, 1982; Perfetti, Landi, & Oakhill, 2005).

Some studies on reading comprehension compared bilinguals and monolinguals and reported that bilingual's reading comprehension was less efficient than those of their monolingual peers (Bonifacci & Tobia, 2017; Jeon & Yamashita, 2014; Melby-Lervåg & Lervåg, 2014, but see also Lipka and Siegel, 2012). When bilinguals' weaknesses related to reading comprehension are investigated, oral language skills, including vocabulary size, rather than decoding appear to be predictive of the delay in reading comprehension acquisition (Lervåg & Aukrust, 2010; Melby-Lervåg & Lervåg, 2013).

SES and reading comprehension in bilingual children

A less explored factor that may interact with reading comprehension acquisition is socioeconomic status (SES), which is generally indexed by parental education, occupational status, and income. In previous literature, children from low-SES families have been found to lag behind their high-SES peers in measures of vocabulary size, grammar, narrative skills, phonological awareness, and speed of language processing (Hoff, 2013).

In the case of bilingual minority children, the conditions of being bilingual and of low SES are often combined. Examining both aspects may be crucial to fully understand the causal nature (and possible overlap) of their relationship with learning and literacy. However, up till now few studies focused on the differential role of SES and being bilingual on reading and comprehension. In a two-year longitudinal study, Droop and Verhoeven (2003) focused on the comparison between monolingual of high and low SES and bilingual children (from North African background) currently living in the Netherlands. Vocabulary size and morphosyntactic knowledge were measured together with reading decoding and reading comprehension, as well as oral language.

Bilingual language minority children had better decoding skills than Dutch low-SES children, but they underperformed relative to high or low SES Dutch children in measures of reading comprehension and oral language proficiency. For all participants, but particularly for bilingual children, reading comprehension skills were better accounted for by the contribution of oral language proficiency than by decoding skills.

More recently, Calvo and Bialystok (2014) investigated the causal role of both being bilingual and SES. The authors tested 175 6-to 7-year-old children from Canada. Half of the tested children were bilingual and half monolingual. Children's linguistic background was assessed through the Language and Social Background Questionnaire (LSBQ). Verbal reasoning, receptive vocabulary, nonverbal and verbal visual attention were assessed through standardized batteries. Flanker task and frog matrices task were used to, respectively, measure inhibition and visuospatial working memory.

It was found that being bilingual was related to lower performance on language tasks, but SES affected both language skills and cognitive skills. In particular, middle class children outperformed working class children on all measures. SES reduced performances in the working-class group compared to the middle class group. The authors state that a way to interpret this result is that SES might be able to delay progresses that children made in mastering the skills required for the proposed tasks. This effect was equivalent for both monolinguals and bilinguals. With respect to being bilingual, bilingual children obtained lower scores in the receptive vocabulary task but higher scores in the flanker task.

In a recent study Bonifacci, Lombardo, Pedrinazzi, Terracina, and Palladino (2020) examined early literacy acquisition in three groups of second graders: monolinguals with high SES, monolingual with low SES and language minority bilingual children with low SES. Bilingual language minority children performed more poorly than

monolingual peers with comparable low SES in reading comprehension and spelling tasks (as shown by a higher number of phonological errors than both SES groups of monolinguals). This result indicates that spelling and reading comprehension in second graders is affected by being bilingual rather than by SES level. However, bilingual children performed comparably to their monolingual peers with low SES in reading decoding and language-comprehension. As a whole, the results indicate that bilingual children require a longer exposure to academic oral language and a longer school practice to reach monolinguals' reading comprehension level. The reason for this longer time required by bilingual children is that both spelling and reading comprehension are skills acquired later in school, deriving from a consolidation in decoding skills and academic oral language.

Howard, Páez, August, Barr, Kenyon, and Malabonga (2014) explored the role that socioeconomic status (SES), home and school language and literacy practices, and oral vocabulary size play in the development of English reading skills in Latino ELLs (English Language Learners). They also examined how all these factors contribute differentially to reading outcomes for children of different ages. Their study considered three different age groups: study 1 was focused on kindergarten children, study 2 on third-grade children, and study 3 on fifth-grade children.

A key finding from the research was that SES, home and school language and literacy environments differentially influenced English reading at each grade level. For kindergarten students, SES and English and Spanish vocabulary knowledge were statistically significant in predicting English word reading. For third-grade students, only Spanish and English vocabulary knowledge were statistically significant in the prediction of word reading, whereas SES, the addition of school language, and the addition of vocabulary knowledge (both Spanish and English) were statistically significant in the

prediction of English comprehension skills. For fifth-grade students, the additions of home language use, school language use, and vocabulary knowledge were statistically significant in the prediction of English word-reading skills. The additions of home language and English vocabulary knowledge were statistically significant in predicting reading comprehension skills.

At the kindergarten level, without any other constructs in the model, SES significantly predicted English reading-accuracy skill, but this was not the case for third or fifth graders. For reading comprehension, while SES was a significant predictor at the third-grade level, it was not at the fifth-grade level. In all, the effects of SES on reading varied according to grade level and reading construct. This may be related to the effects of instruction. The findings show that more exposure at home and/or at school to English predicts better English reading comprehension outcomes.

In sum, the reviewed research points at a role of both being bilingual and SES in affecting vocabulary and literacy development. Available results are not fully consistent, but apparently being bilingual may negatively affect L2 literacy acquisition, an effect that may be reduced by school attendance. In light of this consideration, it may be useful to examine an older sample of monolingual and bilingual students with more than 3 years of school attendance in order to better understand how being bilingual and SES levels may interact with the school exposure increase. Previous studies as the one by Howard and colleagues (2014) may suggest that the role of SES decreases as schooling progresses, but according to Bonifacci et al. (2020) and Droop and Verhoeven (2003), this depends on which specific literacy acquisition is examined. Reading comprehension and spelling in particular appeared to be longer affected by SES.

Aims of the study

In our study we focused on fourth to sixth grade students to investigate reading comprehension in both bilingual and monolingual students. At their academic stage, students are expected to have developed efficient decoding skills and therefore reading experience should have promoted reading comprehension together with vocabulary knowledge. Our aim was to analyze the relationship between vocabulary knowledge and reading comprehension, controlling for SES and being bilingual, in a sample of monolinguals and minority language bilingual children.

Our first hypothesis was that the socioeconomic status of the children's families would play a decisive role in the unfolding of both vocabulary and reading comprehension skills. In contrast, we hypothesized that the influence of being bilingual on vocabulary size and level of comprehension would become less prominent at this school level, in light of the knowledge that the minority language bilingual children already acquired through their school and social context.

As a second hypothesis, we proposed that at the same time the interaction of vocabulary size and SES might be affected by the level of reading comprehension, and, similarly, that the interaction of the level of comprehension and SES would depend on vocabulary size. In other words, low SES would be associated with smaller vocabulary size or lower level of comprehension, but each of these two latter variables might reciprocally explain a portion of the causal relation from SES to the other variable.

Method

Participants

The sample of the present study included 120 two children (53.28% male; 46.72% female), between 9 and 11 years old (M. age: 10.12 years; DS: 0.51), attending public primary and secondary school in Lombardia, a northern Italian region. Of the participants, 54.1% were Italian monolinguals, exposed both at home and at school to Italian. About 45.9% of the children were minority language bilingual children who had been exposed to an L1 other than Italian (L2) within the family context from birth, and to the Italian language at school (30% of the sample speaking Arabic, 27% speaking Albanian, 10% speaking Chinese, 5% speaking Romanian, 4% speaking Brazilian Portuguese, 3% speaking Polish, 1% speaking Turkish and 13% speaking other dialects).

As assessed through the administration of the Italian version of parental questionnaire ALDeQ (Bialystok *et al*., 2016), 9.02 % of bilinguals were born in a foreign country. None of the participants had received a clinical diagnosis of neurodevelopmental disorders. Seven children were excluded from the sample for learning disabilities and other neurodevelopmental disorders, according to Public Health Service.

The Socioeconomic familiar status was evaluated through the Hollingshead Four Factor Index of Social Status (Hollingshead, 1975, see below for a detailed description) and, in terms of the five ranges on which scores can be divided, we found that 19.9 % of participants scored a Low status (8-19 pt.), 26% Medium Low status (20-29 pt.), 18.84% Medium status (30-39 pt.), 22.39 % Medium High status (40-55 pt.), and 12.84 % High status (55-56 pt.). The mean score of participants' SES was 34.89 points (DS: 12.27).

All the parents' participants had to previously sign the informed consent and an oral consent was also obtained by children.

Materials

Socio-economic level (SES).

The Hollingshead Four Factor Index of Social Status (Hollingshead, 1975), adapted to the Italian culture, was applied. For the purposes of the present study, indexes of educational level (EL) and occupation (O) were used. A score from 1 to 7 was applied to the parents' educational level and a score between 1 and 9 to the parents' occupation. SES scores for fathers and mothers were calculated and a compound SES score for children was derived from the sum of the values. Scores between 0 and 39 were classified as low-medium, and scores above 40 as medium-high or high.

Verbal meaning and Primary Mental Abilities (PMA).

The first task that was administered is part of the Italian adaptation of PMA battery (Thurstone & Thurstone, 1981), created with the aim of detecting the specific constitutive factors of intelligence. In the present research, the knowledge of the Verbal Meanings was analyzed through the test "Parole". In particular, this task measures the breadth of vocabulary of the participants, which means the number of words they know.

The PMA test was administered collectively to the class by experimenters.

Breaks were allowed if the child showed signs of fatigue. In this task, children were required to recognize the correct synonym of a word item, written in capital letters, from among four alternatives. The test consists of 30 items and must be completed by the participants in seven minutes (Cronbach's $\alpha = .95$, $r = .92$, as reported in the manual of the test).

MT-3-Clinical Trials.

The second collective test that was administered to the class is part of the MT battery (Cornoldi & Carretti, 2016). This tool is used to assess reading comprehension skills, through a passage reading and multiple choice inferential questions. It identifies difficulties in reading and understanding a text. A booklet was given to every participant. The difficulty of the reading passages was calibrated according to the age of the children.

The booklet contained two different reading passages, a news part and a factual knowledge part. At the end of each part, participants were asked to answer to 12 multiple choice inferential questions. An example of inferential question for the news section is the following: "The two peasants live… A) in two different countries; B) close to each other; C) opposite each other; D) separated by large boulders" (example from "the children bridge", for the IV grade primary school). An example of inferential question for the factual knowledge section is the following: "Who deserves credits for presenting the potato to Louis XVI? A) a European explorer; B) an American explorer; C) a Prussian army officer; D) a French army officer" (example from "the potato discovery", for the IV grade primary school).The class has about 30 minutes to read each passage and answer the related questions. The child is free to choose the best reading strategy (aloud or silent reading) and the passage remains available during answering phase. ("Il ponte dei bambini": Cronbach's $\alpha = .67$, $r = .52$; "La scoperta della patata": Cronbach's $\alpha = .67$, $r = .67$; "Salvataggio della nave in Antartide": Cronbach's $\alpha = .61$, $r = .55$; "Il lago Aral": Cronbach's $\alpha = .73$, $r = .67$, as reported in the manual of the test).

Forward and backward digit-span

This test is part of the Neuropsychological Assessment Battery for developmental age (BVN 5-11, Bisiacchi & Cendron, 2005), which investigates the main cognitive functions (visual perception, memory, language, praxis, attention, higher executive functions, reading, calculation and writing) in children between 5 and 11 years. It allows an analysis of the development of cognitive functions in normal conditions and an identification of any developmental and/ or acquired pathology.

The forward and backward digit-span tests consist of the immediate, direct and inverse serial recall of a string of numbers, previously read by the experimenter and increasing in length. The numbers are presented by the experimenter at the rate of one

number per second. During the backward span, the numerical strings should be repeated in the reverse order, from the last number to the first number. The experimenter marks the participants' responses. The task is divided into blocks of three strings, all of the same length. The test stops when the participant misses two strings out of three of the same block. Trained psychologists administered forward and backward digit-span individually, in a quiet room in the school (Forward Digit-Span: $\alpha = .83$, $r = .81$; Backward Digit-Span: $\alpha = .90$, $r = .82$, as reported in the manual of the test).

Words and Non-Words Reading Task.

The two reading tasks are part of the battery DDE-2 (Sartori, Job & Tressoldi, 2007). For each of the two tasks, three different lists were presented, consisting of words in the former case and consisting of non-word in the latter. Accuracy (counted in terms of number of mistakes, with a maximum of 1 mistake for every word included) and speed (counted in terms of seconds required to read the three lists) were both considered, obtaining an index for both. The final score consisted of the proportion between the two indexes of accuracy and speed ($\alpha = .78$ and $\alpha = .84$, $r = .94$ and $r = .85$, as reported in the manual of the test).

Procedures.

The study was carried out according to the ethical guidelines of the Italian Association of Psychology (AIP) and approved by the Ethics Committee of the Psychology section of University of Pavia.

The experiment started with parents signing the informed consent to allow their children involvement in the study. Furthermore, they performed the Hollingshead Four Factor Index of social status (Hollingshead, 1975) and filled in the ALDeQ parental questionnaire (Bonifacci *et al.*, 2016). The main tasks were administered next to the

children during school hours, within the school's premises, and in accordance with the school schedule.

Children participated in two meetings, one collective and one individual. Teachers took part in the collective meeting, during which the vocabulary size PMA task and the reading comprehension MT3-Cliniche were administered, together lasting about 45 minutes. A 5-minutes break was carried out between the two tasks. During the break, the experimenters were available for answering children's questions.

With respect to the individual meeting, children were called one by one in a separated room, where they performed the two Digit-Span Tasks of about 10 minutes each. Fatigue and perceived difficulty were checked along the whole experiment.

At the end of the experimental phase, every class that took part in the study was involved in a meeting, together with the teachers, in order to explain the main purposes of the research. Furthermore, at the end of the study, the mean of the results obtained by each class involved were reported to teachers and discussed with them, to give them feedbacks regarding the project.

Results

Descriptive statistics.

Data analysis was done using RStudio 1.0.153 (2020). We first report some descriptive statistics regarding the distributions of our main three measures. As can be seen in Figures 2.1-2.3, the results in our three main measures of SES, vocabulary size (as measured by PMA) and reading comprehension (as measured by MT3-Cliniche) followed a normal distribution.

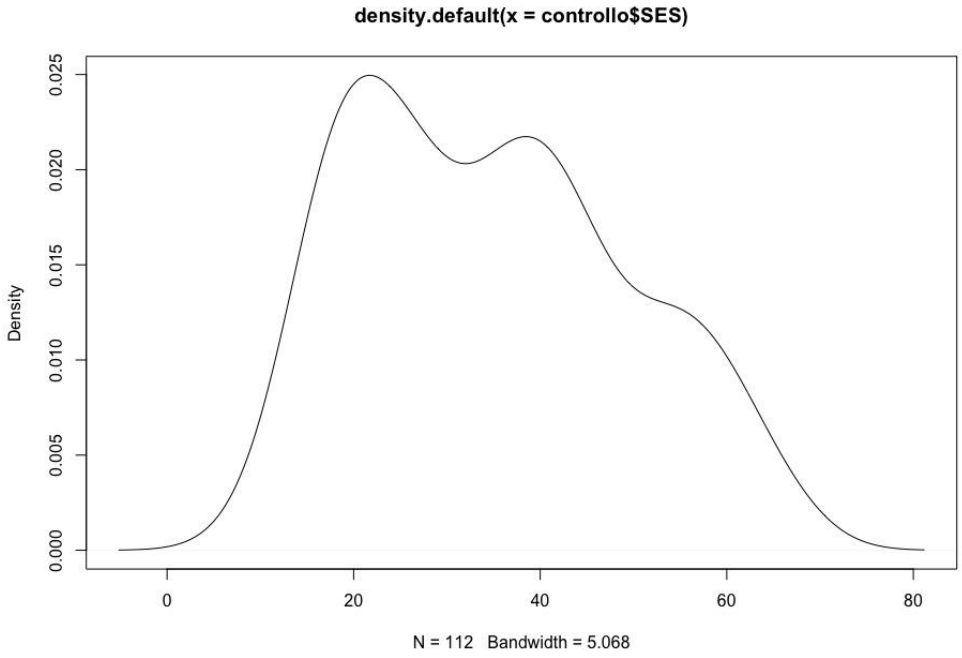


Figure 2.1. SES Density Plot

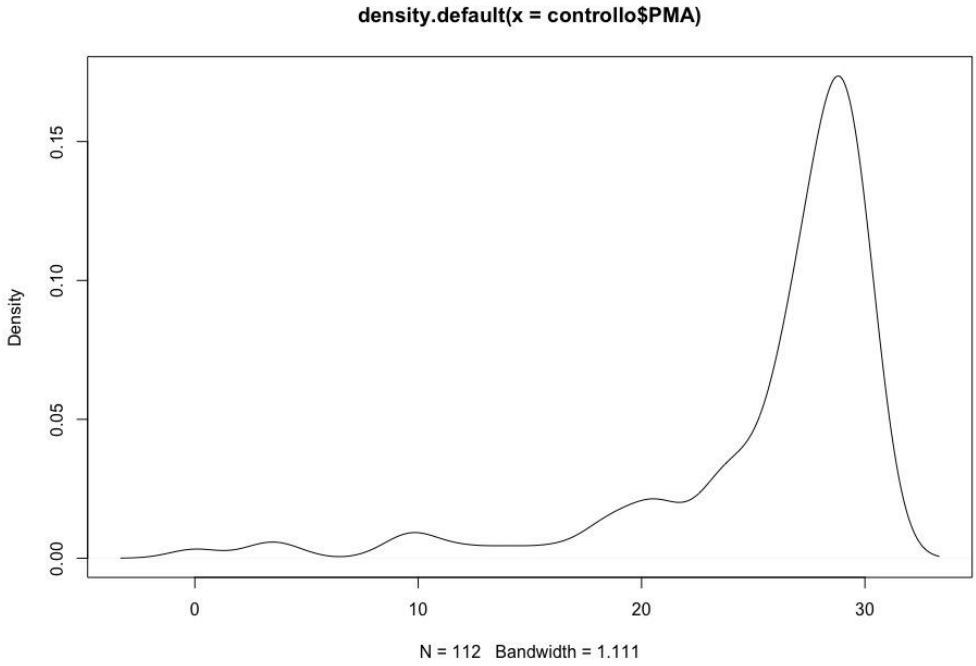


Figure 2.2. Vocabulary Density Plot

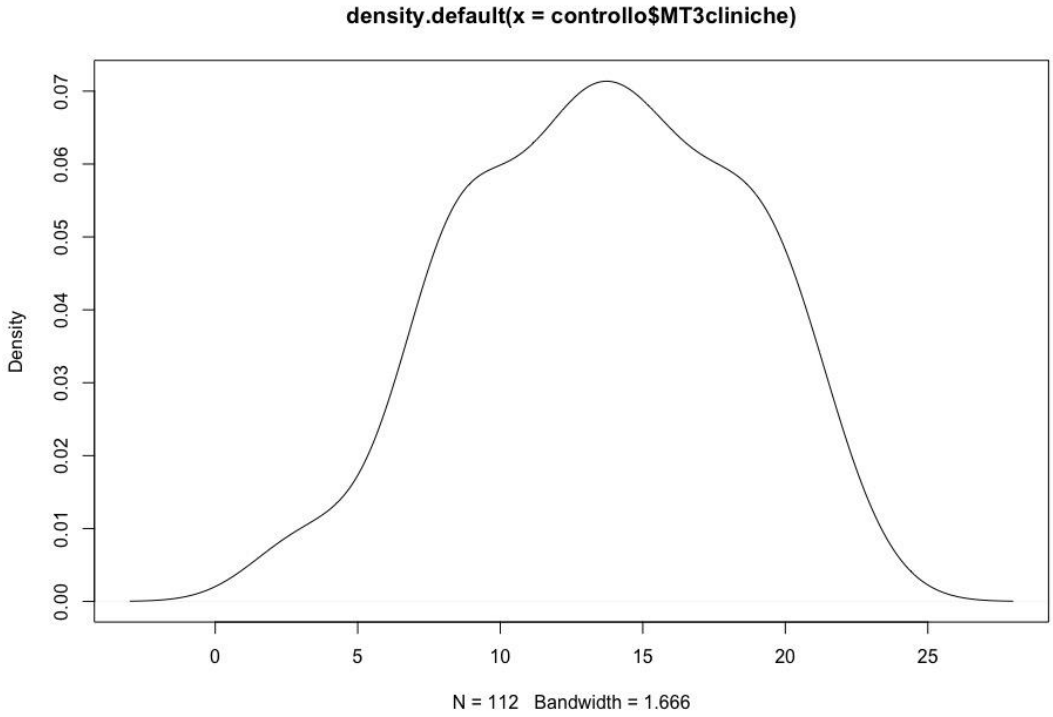


Figure 2.3. Comprehension Density Plot

Table 2.1 reports the correlation matrix between the above mentioned three variables and their significance. All variables were significantly correlated mutually. Relations were further clarified by the following additional analyses.

	Vocabulary Size	Reading comprehension	SES
Vocabulary Size	X	0.43*	0.30*
Reading comprehension	0.43*	X	0.51*
SES	0.30*	0.51*	X

Table 2.1. Between variables correlation matrix

Prior analysis.

As a first step, we considered SES both in its entirety and in its subcomponents (SES of the mother, SES of the father, SES intended as parents' instruction and SES intended as parents' profession).

In particular, the correlation matrix shows that the general SES of the mother (composed by her instruction summed with her profession) correlates highly with the general SES of the father (.55). SES, intended as a total index of instruction. Furthermore, the profession of the two parents strongly correlates with the general SES of the mother (.88) and that of the father (.83).

Thus, the subcomponents of the individual SES (profession and instruction of the mother and of the father) were highly intercorrelated, both when examined within a single gender group (Instruction of the mother correlated .97 with profession of the mother; instruction of the father correlated .70 with profession of the father) and when examined between the two gender groups (instruction of the mother correlated .50 with instruction of the father and profession of the mother correlated .52 with profession of the father). General SES correlated strongly with each of the above mentioned single indexes.

In light of these results, we decided to treat SES as a unitary index, as in the original methodology suggested by Hollingshead (1975).

Main analysis.

Hierarchical Regression

First, we tested the effect of age on our main variables of interest. Its contribution proved not to be significant (for the Age effect on vocabulary size $p = .06$; for Age on reading

comprehension $p = .30$), so we decided not to include it in the main hierarchical regression analysis.

Next, we tested how reading abilities for words and pseudowords affected vocabulary size and reading comprehension, using linear regression. Both independent variables were insignificant in affecting vocabulary size or reading comprehension: in particular, in the regression analysis with vocabulary size as dependent variable, reading ability for words showed a $p = .068$ and reading ability for pseudowords showed a $p = .426$; in the regression with reading comprehension as dependent variable, reading ability for words showed a $p = .109$ and reading ability for pseudowords showed a $p = .364$.

These results indicate that reading abilities did no longer significantly affected vocabulary size or reading comprehension in this age group. We therefore excluded reading ability as an independent variable from our hierarchical regressions.
We included being bilingual as a dichotomous control variable coded as 0 for monolingual participants and 1 for bilingual participants; and we also included Digit-Span treated as an overall index by averaging the Forward Digit-Span and Backward Digit-Span scores. We incorporated these two control variables in the models, because both proved to play a role in isolation. In particular, linear regressions showed that being bilingual is significant in affecting reading comprehension scores ($p = .003$), while span shows a strong trend towards significance in both dependent variables ($p = .054$ for vocabulary scores and $p = .058$ for reading comprehension scores).

We ran two hierarchical regressions, one with vocabulary size as a dependent variable (expressed by vocabulary scorings) and one with reading comprehension as a dependent variable (expressed by reading comprehension scorings) .

First, we performed the regression analysis with PMA vocabulary size task as a dependent variable in different steps. We only considered the intercept at step 1 ($p <$

.001***). At step 2, we added SES as our main predictor ($p < .001$***). At step 3, we added being bilingual as a first control variable (SES remained significant, with a p = .002**, while being bilingual showed no effect, with a $p = .718$). Finally, at step 4 we added Span, which only showed a trend towards significance ($p = .062$); SES remained significant ($p < .001$***)

Second, we performed a similar multi-step regression analysis with the MT reading comprehension task as a dependent variable. The intercept was involved in step 1 ($p < .001$***). At step 2 we added SES as our main predictor ($p < .001$***. At step 3, we added being bilingual as a first control variable (SES remained significant, with a $p <$ .001***, while being bilingual showed no effect, with a $p = .451$). At step 4, we again added Span, which only led to a trend towards significance ($p = .083$); SES maintained its significance ($p < .001$***).[1]

For both regression analyses, one may notice that SES was the only significant variable and that it maintained its significance across the various analysis steps.

Mediation analysis.

Furthermore, we wanted to verify if vocabulary size could at least partially account for the relationship between socioeconomical status and reading comprehension. In order to do this, we ran a mediation analysis. At step 1, we tested the total effect of the independent variable (SES) in the dependent variable (reading comprehension). This first step proved to be significant ($p < .001$ ***).

At step 2, we found that the effect of the independent variable (SES) on the mediator (vocabulary size) was significant ($p < .001$ ***).

At step 3, we found that the effect of the mediator (vocabulary size) on the dependent variable (reading comprehension), taking into account the independent

[1] Trends toward significance were included according to the criteria reported by the R studio outputs

variable (SES), was significant ($p < .001$ ***). The significance found in step 2 makes it clear that there is a mediation of vocabulary in the relation between SES and reading comprehension.

As a further confirmation of this finding, we ran the causal mediation analysis at step 4 using the R package "mediation". This package returns values that are quasi-Bayesian and allow us to observe both the frequentist indexes of *p* values and the estimates, with confidence intervals.

Here we report *p* values for the four indexes obtained: ACME, which tests the total effect of the independent variable on the dependent variable and coincides with the sum of step 2 and step 3 (Estimate 95 % = .030, CI Lower = .009, CI upper = .06, $p <$ .001***); ADE, which accounts for the direct effect of the independent variable on the dependent variable and coincides with step 3 (Estimate 95 % = .144, CI Lower = .089, CI upper = .20, $p < .001$***); the total effect that tests the combination of the direct and indirect effects and that coincides with step 1 (Estimate 95 % = .174, CI Lower = .116, CI upper = .23, $p < .001$***); to conclude, the proportion mediated, which explains how much of the causal effect of the independent variable on the dependent variable can be attributed to the mediator, in terms of shared variance (Estimate 95 % = .168, CI Lower = .052, CI Upper = .34, $p < .001$ ***). All the four values showed to be significant, as indicated by *p* values. (See Figure 2.4 and Figure 2.5)

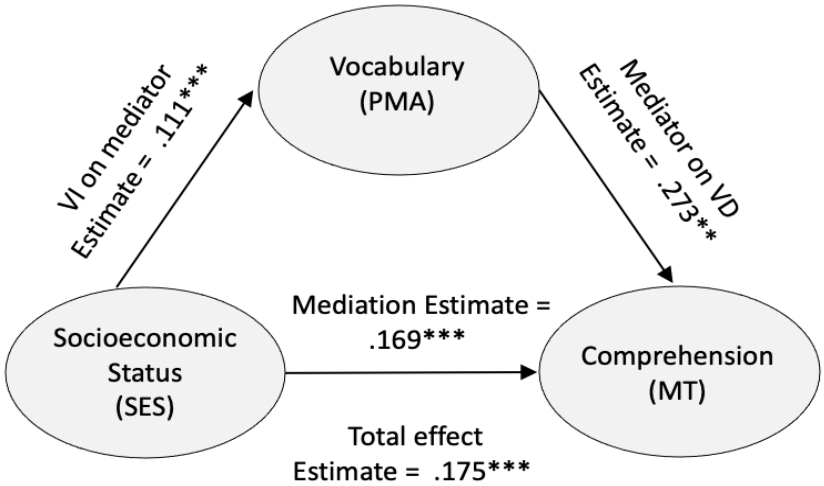


Figure 2.4. Mediation model with PMA as mediator. Along the bottom line is represented step 1; along the left arrow is represented step2; along the right arrow is represented step 3; in the center of the figure is represented the mediation effect obtained in step 4.

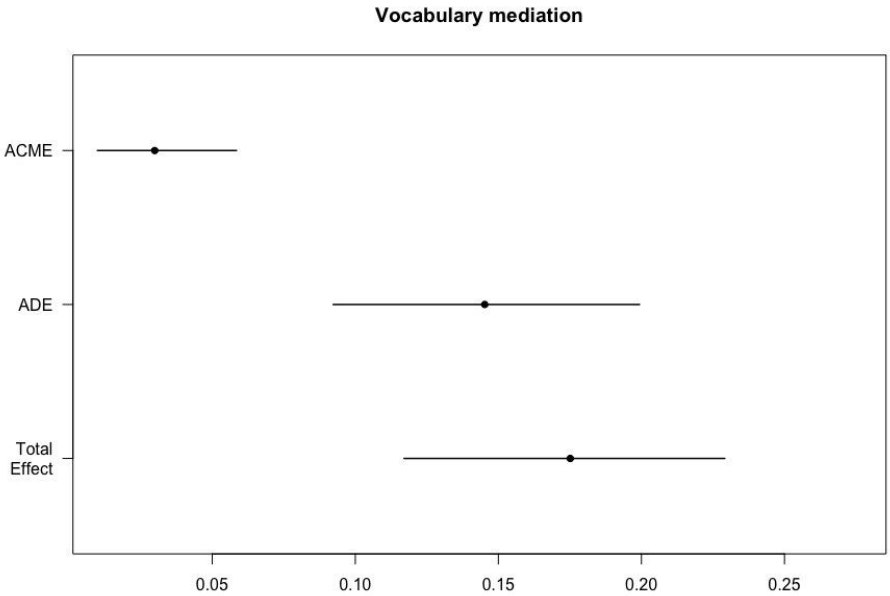


Figure 2.5. Step 4 of the mediation analysis with PMA as mediator. The vertical axis lists indirect, direct, and total effects, while the horizontal axis indicates the respective magnitudes: each central point represents the averages of the distribution, while the horizontal lines represent the amplitude of the distributions, fixated at the 95 % percentiles.

We also decided to test for effects in the opposite direction to the previously exposed one, and to disentangle the effect of reading comprehension on vocabulary skills. To this end, we ran a regression with vocabulary size as dependent variable and reading comprehension as independent variable, adding SES as a covariate. Reading comprehension played a significant role on vocabulary skills, with $p < .001$, while SES was not significant in this direction ($p = .321$). (See Figure 2.6)

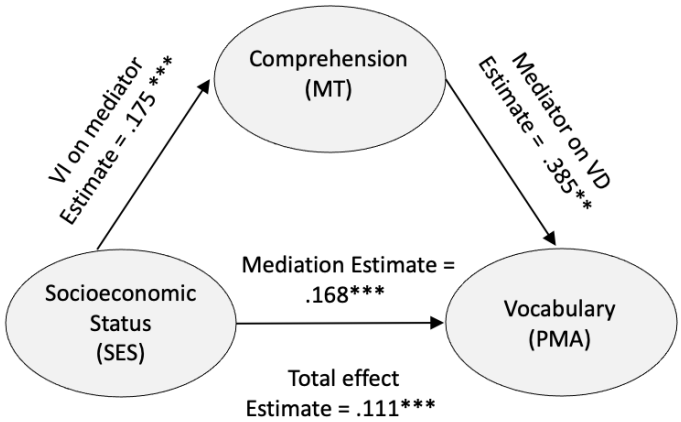


Figure 2.6. Mediation model with MT3 Cliniche as mediator. Along the bottom line is represented step 1; along the left arrow is represented step; along the right arrow is represented step 3; in the center of the figure is represented the mediation effect obtained in step 4.

We also repeated the mediation analysis with reading comprehension as a mediator.

At step 1, we tested the total effect of the independent variable (SES) in the dependent variable (vocabulary size). This first step proved to be significant ($p < .001$ ***). At step 2, we tested the effect of SES on the mediator (reading comprehension) and found it to be significant ($p < .001$ ***). At step 3, we tested the effect of the mediator (reading comprehension) on the dependent variable (vocabulary size), taking into account the independent variable (SES). Reading comprehension resulted in a significant effect (p

= .001 **). The significance found in step 2 makes it clear that there is a mediation of comprehension in the relation between SES and vocabulary size. At step 4, we ran the causal mediation analysis.

Here we report *p* values for the four indexes obtained: ACME, that tests the total effect of the independent variable on the dependent variable. It coincides with the sum of step 2 and step 3 and it was significant (Estimate 95 % = .068, CI Lower = .026, CI upper = .12, *p* = .002**); ADE, that accounts for the direct effect of the independent variable on the dependent variable. It coincides with step 3 and proved not to be significant (Estimate 95 % = .044, CI Lower = -.026, CI upper = .11, *p* = .228); the total effect that tests the combination of the direct and indirect effects. It coincides with step 1 and was significant (Estimate 95 % = .112, CI Lower = .050, CI upper = .18, *p* < 001***) and the proportional mediated, that explains how much of the effect of the independent variable on the dependent variable can be attributed to the mediator and it was significant (Estimate 95 % = .615, CI Lower = .235, CI Upper = 1.42, *p* = .002**). (See Figure 2.7)

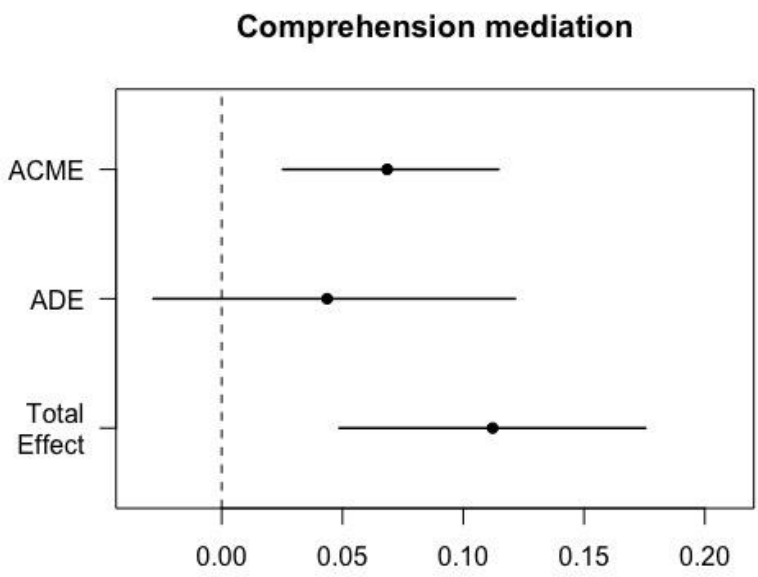


Figure 2.7. Step 4 of the mediation analysis with MT3 Cliniche as mediator. The vertical axis lists indirect, direct, and total effects, while the horizontal axis indicates the

respective magnitudes: each central point represents the averages of the distribution, while the horizontal lines represent the amplitude of the distributions, fixated at the 95 % percentiles.

Discussion

In most of the European schools more and more students are bilinguals (about 10 % of the students according to Invalsi open data, updated to 2021). In particular, most of them can be defined as minority language bilinguals, i.e., persons using an L1 language that is spoken by the minority of the population of a territory. In general, a language spoken by a percentage of less than 50% of the population is defined as minority (Bonifacci, 2016). Minority language bilingualism is often associated with low parental income (e.g., Hoff, 2013), but, as has been argued, it is important to consider the influence of socioeconomic status on being bilingual (see, e.g., Bonifacci et al., 2020; Morton & Harper, 2007; both studies focusing on developmental trajectories of bilingual literacy).

The present study aimed at better understanding the reciprocal relation between vocabulary size and reading comprehension when being bilingual and SES are taken into account. We investigated this issue for a representative sample of children with few years of literacy school learning. The relation between vocabulary size and reading comprehension was found to be bidirectional, in support of the available literature (e.g., Quinn *et al.*, 2020). The reciprocal relations were further affected by SES and not by being bilingual when students had already received 3-5 years of education in their L2.

Our study had two aims. The first one was to clarify which variables accounted, respectively, for vocabulary skills and reading comprehension skills between SES, being bilingual, and working memory. Only SES played a significant role in relation to both vocabulary skills and reading comprehension skills. This result seems to be opposite to that by Bonifacci *et al.* (2020). When the authors measured level of reading comprehension and spelling in second grade children, divided in a monolingual and a

bilingual group, being bilingual negatively impacted participant performance, even when SES was included in the analysis. The finding that being bilingual did not play a role in our sample, but did in younger children' literacy skills (Bonifacci *et al.,* 2020) might be interpreted as evidence that schooling can soften the effect of being bilingual on vocabulary size. An increase in school attendance may compensate for the effects of being bilingual, because vocabulary size measurements are known to be poorer in each of the two languages of a bilingual child in comparison with monolingual peers, at least during early literacy (see also Bialystok, 2009). However, SES could account for difficulties in the literacy domain more than being bilingual. These two aspects, being bilingual and low socioeconomic status, were significantly and negatively correlated in our sample. Bilingual children belonged to families with lower parental SES compared to monolingual children, consistent with the general literature (as in Lonigan *et al.*, 2013). Nevertheless, only SES was able to account for the differences in vocabulary size and reading comprehension skills independently from being bilingual. Therefore, we conclude that being bilingual is no longer affecting literacy after few years in a school context. In contrast, the familiar socioeconomic status intended as an average of instruction and profession of both parents would still play a significant role (e.g., Dolean *et al*., 2019, Fikrat-Wevers, van Steensel & Arends, 2021).

The second aim of the present study was to better understand the relation between vocabulary size and reading comprehension and the role that each of them plays in relation to SES. We ran two mediation analyses to assess these aspects, considering vocabulary size mediated reading comprehension skills in the causal relation between SES and reading comprehension. As an explanation we propose that the breadth and depth of one's vocabulary size correlate with one's chances to correctly extract the

meaning of a text and the messages it is intended to convey (e.g., Droop & Verhoeven, 2003; Howard, 2014).

Remarkably, we found that also reading comprehension mediated the casual relation between SES and vocabulary scores. This aspect might imply that, at this age level, vocabulary skills cannot be considered as detached from reading comprehension. As the present results indicate, vocabulary size and level of reading comprehension exert a mutual influence, while they both are affected by SES (e.g., Colé *et* al., 2018; Lervang *et al*., 2019).

In sum, SES was found to play a fundamental role in affecting literacy skills, regardless of being bilingual. However, the skills a child acquires in the vocabulary size and in the reading comprehension domain might both mitigate this effect. These results are relevant from an applied perspective, as they shed light on the importance of the school system in the developing individual.

A limitation of our study lies in the ignored effects of the L1 of our bilinguals and in the variability of the languages that we included in a unique category of "minority languages". Furthermore, we were not able to apply or adapt the SES questionnaire and the bilingual questionnaire to the parents' L1, because too many languages were involved.

As a future development of our research, our sample could be expanded in order to allow a division of our group of minority language bilingual children into homogeneous subgroups of bilinguals with specific mother tongues. This first proposal would allow us to bring the lens closer to the subgroups of minority language bilinguals A second possibility would be to include an additional bilingual participant group, composed by majority language children with the same characteristics of the minority language ones but speaking "majority languages" such as English and French. Both types

of future research would deepen our understanding of the role of SES in various types of bilinguals and pay greater attention to the linguistic and cultural specificities of our sample. This second proposal would help to better define the characteristics that can make the macro-group of minority language bilinguals more homogeneous as opposed to that of majority language bilinguals.

Yet another possibility would be to include oral comprehension skills as a variable, since it might have a different impact and/or be differently affected by our main variables in comparison with reading comprehension (e.g., Ouellette & Show, 2014), since a better oral comprehension might predict better results on reading comprehension, too but might also be explained by vocabulary and word knowledge at the level of meaning (e.g., Lervåg, Hulme, & Melby-Lervåg, 2018).

Conclusion

The analyses of our results clearly demonstrate that socioeconomic status (SES) plays a significant role in the unfolding of vocabulary size and reading comprehension skills in children aged from 9 to 11. About half of our participant sample was composed by minority language bilingual children, with the other half being composed by a monolingual control group of children with comparable age and schooling. Only SES was found to affect vocabulary size and comprehension skills, while being bilingual for our age group no longer showed the influence observed for younger children (e.g., Bonifacci *et al.*, 2020). This result appears true despite the high correlation that exists between minority language bilingualism and lower socioeconomic levels, already addressed in the literature (e.g., Hoff, 2013). As a further step, our study made clear that vocabulary size is a determinant of the unfolding of comprehension skills in this age group. This is similar to for reading comprehension, which also proved to be determinant in enriching a child's vocabulary size. In summary, all our three main variables (namely SES, vocabulary size,

and reading comprehension) are interacting determiners of literacy development at this stage, irrespective of the child is bilingual or not.

These results become even more explanatory if we focus on the category of minority language bilingual children. Minority language bilingualism is socially perceived as profoundly different from what we are used to think when we talk about "bilingualism". This aspect might depend on the fact that the largest part of the languages spoken by our sample fall into the definition of "subtractive bilingualism".

With this definition we refer to those cases in which one of the languages spoken by a bilingual is a minority language and does not offer additional benefits, at a social, relational, working/school level (e.g., Morandi, 2014). In this cases, the social context of the other dominant language spoken (the L2 of our children) does not offer a real chance to give relevance and to return specificity to the vast multiplicity of languages that fall under the general definition of "minority language bilingualism" (e.g., Landry & Allard, 1993). This consideration shed further light on our result, according to which the variable "bilingualism" gradually loses relevance, while the socioeconomic status of the child's family remains determinant. This latter aspect might be interpreted in light of the different possibilities of access to cultural and practical resources (e.g., remedial courses, tools present at home ...) which can affect the learning process (Dietrichson *et al.*, 2017).

The interaction between socioeconomic status, vocabulary size and reading comprehension seems to be circular: each of these variables plays a role on the others. In light of this interaction, the school might break the circle, since the achievement of a good knowledge in our variables of interest can affect literacy, despite SES.

Acknowledgements, We thank Borrello Mariarita and Squaranti Luca for their precious help in the data collection phase. To conclude, we thank the Head Teachers of the Istituto Comprensivo

Acerbi of Pavia where we collected data, Prof. Rimini Luisa and Prof. Bassi Elena and all the teachers of the classes that took part in our study for their involvement and support.

References

August, D. e Shanahan, T. (Eds.). (2007). *Developing Reading and Writing in Second-Language Learners: Lessons from the Report of the National Literacy Panel on Language-Minority Children and Youth Published by Routledge for the American Association of Colleges for Teacher Education*. Routledge.

Beck, I. L., McKeown, M. G. e Kucan, L. (2002). Bringing words to life: Robust vocabulary instruction. New York: Guilford.

Beck, I. L., Perfetti, C. A. e McKeown, M. G. (1982). Effects of long-term vocabulary instruction on lexical access and reading comprehension. *Journal of educational psychology*, *74*(4), 506.

Bialystok, E. (2009). Bilingualism: The good, the bad, and the indifferent. *Bilingualism: Language and Cognition*, *12*(1), 3–11.

Bonifacci, P., Lombardo, G., Pedrinazzi, J., Terracina, F. e Palladino, P. (2020). Literacy skills in bilinguals and monolinguals with different SES. *Reading & Writing Quarterly*, *36*(3), 243-259.

Bonifacci, P., Mari, R., Gabbianelli, L., Ferraguti, E., Montanari, F., Burani, F., & Porrelli, M. (2016). Sequential bilingualism and specific language impairment: The Italian version of ALDeQ parental questionnaire. *BPA-Applied Psychology Bulletin* (Bollettino di Psicologia Applicata), *64*, 50–64.

Bonifacci, P. e Tobia, V. (2017). The simple view of reading in bilingual language-minority children acquiring a highly transparent second language. *Scientific Studies of Reading*, *21*(2), 109-119.

Calvo, A. e Bialystok, E. (2014). Independent effects of bilingualism and socioeconomic status on language ability and executive functioning. *Cognition*, *130*(3), 278-288.

Colé, P., Cavalli, E., Duncan, L. G., Theurel, A., Gentaz, E., Sprenger-Charolles, L., & El-Ahmadi, A. (2018). What Is the Influence of Morphological Knowledge in the Early Stages of Reading Acquisition Among Low SES Children? A Graphical Modeling Approach. *Frontiers in Psychology*, *9*.

Cornoldi, C., & Carretti, B. (2016). Prove MT-3-Clinica: La valutazione delle abilita di lettura e comprensione (MT-3 tests for clinical work: Assessing reading and comprehension abilities). Firenze, Italy: Giunti EDU

Dietrichson, J., Bøg, M., Filges, T., & Klint Jørgensen, A. M. (2017). Academic interventions for elementary and middle school students with low socioeconomic status: A systematic review and meta-analysis. *Review of Educational Research, 87*(2), 243-282

Dolean, D., Melby-Lervåg, M., Tincas, I., Damsa, C., & Lervåg, A. (2019). Achievement gap: Socioeconomic status affects reading development beyond language and cognition in children facing poverty. *Learning and Instruction*, *63*, 101218.

Droop, M., e Verhoeven, L. (2003). Language proficiency and reading ability in first-and second-language learners. *Reading research quarterly*, *38*(1), 78-103.

Fikrat-Wevers, S., van Steensel, R., & Arends, L. (2021). Effects of Family Literacy Programs on the Emergent Literacy Skills of Children From Low-SES Families: A Meta-Analysis. *Review of Educational Research*, 003465432199807.

Gough, P. B. e Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remedial and special education*, *7*(1), 6-10.

Grosjean, F. (1992). Another view of bilingualism. In *Advances in psychology* (Vol. 83, pp. 51-62). North-Holland.

Gernsbacher, M. A. (1997). Two decades of structure building. *Discourse processes*, *23*(3), 265-304.

Hoff, E. (2013). Interpreting the early language trajectories of children from low-SES and language minority homes: implications for closing achievement gaps. *Developmental psychology*, *49*(1), 4.

Hollingshead, A. B. (1975). *Four factor index of social status. Working paper.* New Haven, CT: Yale University Press.

Hoover, W. A. e Gough, P. B. (1990). The simple view of reading. *Reading and writing*, *2*(2), 127-160.

Howard, E. R., Páez, M. M., August, D. L., Barr, C. D., Kenyon, D. e Malabonga, V. (2014). The importance of SES, home and school language and literacy practices, and oral vocabulary in bilingual children's English reading development. *Bilingual Research Journal*, *37*(2), 120-141.

Jeon, E. H. e Yamashita, J. (2014). L2 reading comprehension and its correlates: A meta-analysis. *Language learning*, *64*(1), 160-212.

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge university press.

Landry, R., & Allard, R. (1993). Beyond Socially Naive Bilingual Education: The Effects of Schooling and Ethnolinguistic Vitality on Additive and Subtractive Bilingualism.

Lecce, S., Zocchi, S., Pagnin, A., Palladino, P., & Taumoepeau, M. (2010). Reading minds: The Relation Between Children's Mental State Knowledge and Their Metaknowledge About Reading. *Child Development, 81*, 1876–1893.

Lervåg, A. e Aukrust, V. G. (2010). Vocabulary knowledge is a critical determinant of the difference in reading comprehension growth between first and second language learners. *Journal of Child Psychology and Psychiatry*, *51*(5), 612-620.

Lervåg, A., Hulme, C., & Melby-Lervåg, M. (2018). Unpicking the developmental relationship between oral language skills and reading comprehension: It's simple, but complex. *Child Development, 89*(5), 1821-1838.

Lervåg, A., Dolean, D., Tincas, I., & Melby-Lervåg, M. (2019). Socioeconomic background, nonverbal IQ and school absence affects the development of vocabulary and reading comprehension in children living in severe poverty. *Developmental Science*, *22*(5).

Lipka, O. e Siegel, L. S. (2012). The development of reading comprehension skills in children learning English as a second language. Reading and Writing, 25(8), 1873-1898.

Lonigan, C. J., Farver, J. M., Nakamoto, J., & Eppe, S. (2013). Developmental trajectories of preschool early literacy skills: A comparison of language-minority and monolingual-English children. *Developmental Psychology, 49*(10), 1943–1957.

Melby-Lervåg, M. e Lervåg, A. (2014). Reading comprehension and its underlying components in second-language learners: A meta-analysis of studies comparing first-and second-language learners. *Psychological bulletin*, *140*(2), 409.

Moradi, H. (2014). An investigation through different types of bilinguals and bilingualism. *International Journal of Humanities & Social Science Studies, 1*(2), 147-154.

Morton, J. B., & Harper, S. N. (2007). What did Simon say? Revisiting the bilingual advantage. *Developmental Science*, *10*(6), 719–726.

Oakhill, J. (2020). Four decades of research into children's reading comprehension: A personal review. *Discourse Processes*, *57*(5-6), 402-419.

Ouellette, G. & Shaw, E. (2014). Oral vocabulary and reading comprehension: An intricate affair. *L'Année psychologique, 114*, 623-645.

Perfetti, C. e Adlof, S. M. (2012). Reading comprehension: A conceptual framework from word meaning to text meaning. *Measuring up: Advances in how we assess reading ability*, 3-20.

Perfetti, C. A. e Hart, L. (2001). The lexical basis of comprehension skill.

Perfetti, C. A., Landi, N. e Oakhill, J. (2005). The acquisition of reading comprehension skill.

Quinn, J. M., Wagner, R. K., Petscher, Y., Roberts, G., Menzel, A. J., & Schatschneider, C. (2020). Differential codevelopment of vocabulary knowledge and reading comprehension for students with and without learning disabilities. *Journal of Educational Psychology, 112*(3), 608–627.

RStudio Team (2020). RStudio: Integrated Development for R. RStudio, PBC, Boston, MA URL http://www.rstudio.com/

Sartori, G., Job, R. & Tressoldi, P. E. (2007). *DDE-2 . Batteria per la valutazione della dislessia e della disortografia evolutiva – 2.* Firenze, O. S.

United Nations Educational, Scientific and Cultural Organization (UNESCO). (2018). Global education monitoring report 2019: MIGRATION, displacement and education: Building bridges, not walls.

Verhoeven, L. e Van Leeuwe, J. (2008). Prediction of the development of reading comprehension: A longitudinal study. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, *22*(3), 407-423.

Verhoeven, L., van Leeuwe, J. e Vermeer, A. (2011). Vocabulary growth and reading development across the elementary school years. *Scientific Studies of Reading*, *15*(1), 8-25.

During the previous two chapters we addressed the role of emotionality on verbal false memories' production as a sign of semantic processing in bilingual children, speaking a minority language as L1 and Italian as L2. We addressed that there were no differences between these children and their monolingual peers in terms of processing of emotional words. Thus, our results lead us to consider their semantic processing skills as comparable to the ones of their monolingual peers. Furthermore, we considered and deepened the role of their socioeconomic status, since it turned out to be a determinant variable for the acquisition of linguistic and comprehension skills.

In the next chapter we are going to consider a completely different population of bilinguals. Indeed, they are adults leaving in Bozen, speaking German as L1 and Italian as L2. Our aim is to consider also in this group the role that emotionality of the words have on their verbal false memories' production and consequently on their semantic processing. We decided to include them, since they speak a majority language as L1 that is highly supported and integrated in the social context in which they leave. Thus, we think this group can provide further insights concerning the more general topic on emotional semantics in bilinguals.

CHAPTER 4

The role of emotionality on verbal false memories in German/Italian bilinguals.

In preparation: Cangelosi, M., Cottini, M., & Palladino, P. (2022).
To be submitted to *Bilingualism: Language and Cognition*.

Abstract.

Studies show that words with high emotionality (e.g., 'Pain') are processed more extensively, thus stimulating semantic processing (e.g., Zhang, 2016). To assess how emotionality (arousal and valence) affect the production of verbal false memories, the DRM paradigm for assessing false memories has been extended by Brainerd and colleagues (CEL Lists, 2008) . In the past, the DRM has been frequently applied to monolinguals of different ages, but not yet to bilinguals. Filling this gap can shed light on both the emotional and the semantic processing of bilinguals, which make up two research domains that are still subject to a wide debate. We measured emotional verbal false memories through the CEL Lists on a population of German/Italian bilinguals living in Bozen, a city in Northern Italy. Our results showed that the bilinguals of our sample are less accurate and produce less false memories in comparison with a monolingual control group. With respect to the emotional component, we found that positive valence and low arousal lures significantly reduced the number of false memories in a participant. These results are discussed in the light of the literature and considering different types of bilinguals.

Keywords. Semantics; False Memory; Arousal; Valence; Bilingualism.

Introduction

Words are not sterile. Together with the meaning they convey, they are also bearers of a certain connotation, which we can define in terms of their arousal (i.e., the stronger or weaker emotional activation they elicit) and valence (i.e., negative or positive). As an example, consider the Italian word ‘rabbia’ (meaning ‘anger’ in English). For Italian monolinguals, this word belongs to their native language and is well-established in memory by virtue of its high frequency of use. But “rabbia” also has a strong emotional connotation, being a word with high arousal and negative valence. This emotionality will prioritize the elaboration of "rabbia" if it occurs between other words in a sentence. The urgency and attraction of focus of words with a strong emotional content make them special and stimulate their processing in terms of their semantics. This is a specially so for the meaning aspects that relate to the context in which they are inserted, rather than on the specific and superficial characteristics of "rabbia" (e.g., where "rabbia" is inserted in a list of words, in which color it is written, in which font...). In contrast, for bilinguals who speak German as their first language and Italian as their second language, the word "rabbia" is just the translation of the core meaning "Wut", with which they would express the concept of anger in their first language. It can be questioned how much would remain of the prioritizing and subsequent semantic processing for "rabbia" by the Italian monolingual. Will the Italian word be so deeply rooted in the German/Italian’s bilingual knowledge both in terms of meaning and of emotionality as to be comparable to how the Italian monolingual would process it?

The aim of this study is to investigate how the emotional connotation of words influences verbal processing in bilinguals through the analysis of their verbal false memories. In particular, we want to test if for a group of German/Italian bilinguals the

greater emotionality associated with a term presented in their L2 (Italian) acts as a booster for semantic processing, enhancing false memories for emotional words, similarly to what has already been found in monolingual groups (e.g., Zhang, Gross & Hayne, 2016).

Verbal false memories have been classically measured by means of the Deese-Roediger-McDermott paradigm (DRM paradigm, Deese, 1959; Roediger & McDermott, 1995). In the DRM paradigm a person is presented with a list of words, one at a time, all semantically associated with a theme not present in the list (commonly called "critical lure"). When that person is asked to memorize as many of the given words as possible, it is likely that he or she will recall or recognize also the critical lure, as if it would actually have been present in the list.

This phenomenon has been argued to occur as a consequence of semantic activation spreading during the encoding of words. More specifically, there are two accounts: Fuzzy Trace Theory (FTT; Reyna *et al.,* 2016) and by Spreading Activation Theory (SAT; Roediger, Balota & Watson, 2001). Fuzzy Trace Theory postulates that there are two independent types of processing forming the basis of recollection. The first is verbatim retrieval, through which we elaborate the surface characteristics of a trace to be memorized (e.g. the color of a stimulus or its position among other elements). The second is gist retrieval, through which we elaborate the meaning conveyed by what has been processed, namely its semantics (e.g., Brainerd & Reyna, 2004; Reyna, Corbin, Weldon & Brainerd, 2016; Howe & Derbish, 2010; Otgaar, Howe, Muris & Merckelbach, 2018). Alternatively, Spreading Activation Theory postulates that processing a concept results in spreading activation to related ones, thus inducing false memories when these are “incorrectly” activated semantically as intruders.

Brainerd, Yang, Reyna, and Howe (2008) tested the properties of the DRM paradigm and found that it can be reliably used to measure semantic processing. More in detail, they analyzed 16 dimensions of semantic content to assess their ability to predict interlist variability in false memory production. Fifteen of the sixteen dimensions turned out to be highly correlated with forward or backward association strength. Forward association indexes refer to the strength according to which the critical lure is semantically associated with each of the words present in the list; on the contrary, backward association indexes refer to the strength according to which every word present in the list is semantically associated with the critical lure. Both of these are highly used to build DRM lists.

Concerning the words' emotional content, the vast majority of research studies demonstrated that a higher number of false memories is elicited by high arousal lures relative to neutral ones.

For example, Dehon, Larøi, and Van der Linden (2010) presented participants with DRM lists that mainly consisted of negative words, DRM lists that mainly consisted of positive words, or DRM lists that mainly consisted of neutral words, asking participants to later recall or recognize them. Affective lists (both the negative and positive ones) elicited more false memories for their critical lures than neutral lists, both during the recall and the recognition tasks.

Similarly, using emotional and neutral DRM lists, Sharkawy, Groth, Vetter, Beraldi, and Fast (2008) found that critical lures that were associated with emotional word lists were significantly more often falsely recognized than neutral ones.

With respect to valence, lists that were associated with critical lures with negative valence produced a larger number of false memories. For example, Brainerd, Stein,

Silveira, Rohenkohol, and Reyna (2008) compared the effect of DRM word lists with negative valence against positive and neutral ones. Lists with negative valence were associated with the highest incidence of false memories. According to their view, positive valence would be protective with respect to memory distortion. This conclusion comes from an analysis made on three parameters, namely, Recollection rejection, Phantom Recollection and Similarity judgment. As a whole, they found two main results: the former was that negative valence of the items enhances the familiarity of the semantic contents of the critical lures, while positive valence makes the opposite effect; the latter was that negative valence impairs the ability to use verbatim traces to suppress memory errors, whereas positive valence makes the opposite effect.

Nevertheless, for both valence and arousal, results seem to be dependent on the nature of the task. There largest unanimity has been found when participants are asked to recognize the DRM lists. On the contrary, when they are asked to recall them, results become less clear.

For example, Dehon, Larøi and Van der Linden (2010) found that when participants were asked to recall their DRM lists, Negative and positive critical lures were equally more recalled than the neutral ones.

Brainerd, Yang, Toglia, Reyna and Stahl (2008) validated a pool of DRM lists whose critical lures are divided on the basis of their arousal and of their valence, in a 4 x 4 factorial design (Cornell/Cortland Emotion Lists, CEL Lists, 2008).

CEL Lists have been applied to several monolingual groups, belonging to different ages (Brainerd, Holliday, Reyna, Yang & Toglia, 2010). However, from our analysis of the literature it seems that these lists have not so far been tested on bilinguals.

What do we know about the relation between emotionality and semantics in bilinguals and about bilinguals' emotional processing?

The relation between semantic meaning representation in L1 and emotionality has already been studied through other tasks. It is more complex to study this interaction across languages (e.g., Sianipar, Middelburg & Dijkstra, 2015).

Altarriba and Canary (2004) investigated how the arousal effects word processing in the L1 and in the L2 semantic networks. More in detail, they proposed a semantic priming study both in the L1 and in the L2 of proficient bilingual speakers. They aimed to assess if the arousal congruency of an high arousing target and its preceding prime slowed the performance. They found that congruency facilitated the task, both when the primes were high arousing or moderate arousing. On the contrary, non-arousing primes did not enhanced the performance. As a whole, this result seems to show that arousal congruency acts on the semantic networks of both the L1 and the L2 of proficient bilinguals.

As regard valence, Harris, Gleason and Ayçiçegi (2006) showed that different types of emotional stimuli modulated L2-dominant speakers' skin conductance responses (SCR) to both L1 and L2 verbal stimuli. The emotion effect was stronger in L1 when the stimuli were negative and presented as childhood reprimands. However, in general, emotion effects were also present in SCR to L2 stimuli.

Several other studies investigated if emotionality is reduced in L2. For example, Degner, Doycheva and Wentura (2012) reported a valence priming facilitation effect in the L2 of highly proficient bilingual speakers with high levels of immersion and frequency of L2 use. This result suggests that the emotional associations of L2 representations might indeed be determined by the bilinguals' experiences with L2 (i.e., age of acquisition, levels of L2 exposure / L2 proficiency).

In a previous study (Cangelosi, Bossi, & Palladino, 2021), we investigated the effect of emotionality on verbal false memories using the same emotional DRM pool (CEL Lists, Brainerd *et al.*, 2008) on a sample of minority language bilingual children attending from IV to VI grade and compared with a monolingual control group of their same age. The DRM words were presented in the bilingual children L2, namely Italian. We found that our bilingual sample did not differ in the way in which emotion (intended as a combination of arousal and valence) affected verbal false memories. More in detail, we found that the negative valence-high arousal lures were the ones that were more often falsely remembered in both the groups. SES was considered in our experimental design and it turned out to be significant in affecting verbal false memory. In fact, lower levels of SES corresponded to a lower degree of false memories for semantically related critical lures.

Even if the age considered in our previous study (2021) was different from the age of the sample of the present study, we would like to start from these results in order to consider how emotionality acts on the semantic processing that is involved in verbal false memories in bilingual populations that especially differ in terms of SES and of linguistic status of their first language.

Overview

We found a gap in the literature regarding how false memories', intended as an index of semantic processing, are affected by emotionality of the words in bilinguals' L2.

We chose to propose an adapted to Italian version of the CEL Lists to a sample of German/Italian bilingual adults from Bozen. Bozen is the capital of Trentino Alto-Adige, an autonomous region located in Northern Italy.

The city can be considered as fully bilingual, since both German and Italian are recognized as official languages.

These characteristics make this sample of particular interest as it is a population in which bilingualism is rooted at all levels of society (e.g., schools, work, contexts of daily life) and, therefore, strongly encouraged and supported on a social level.

Indeed, studying how the connotative dimension of semantics develops in this population of bilinguals can shed light on bilingual processing in a domain that lies on the edge between semantics and emotion, and is still strongly debated (e.g., Dong, Gui & Macwhinney, 2005; Altarriba & Basnight-Brown, 2009; Braunstein *et al*., 2012; Khateb, Pegna, Michel & Mouthon, 2016; Francis, 2018; Novitskiy, Myachykov & Shtyrov, 2018).

Does semantics develop so deeply as to be comparable to the one of monolingual peers also in terms of emotionality that meaning can convey? In light of this, is this semantic processing so rooted in the bilingual mind as to cause as many false memories as for the monolinguals, with the same emotional pattern?

Or, on the contrary, does the second language spoken by this kind of bilinguals lack that deep emotionality? Does it consequently protect them against memory distortion?

Aims of the study

The aim of our study was to verify how the emotional connotation of a pool of DRM lists influences false memory, intended as a valid index of semantic processing, in the second language of a sample of German/Italian bilingual adults, in comparison with a group of monolinguals comparable in terms of age and education.

In particular, our interest was to verify the assumption according to which the greater emotional connotation of the actually coded terms (divided factorially on the basis of their arousal and their valence) favors a semantic based (or gist based) processing, to the detriment of a more precise analysis of the specific characteristics of the material that

is required to be memorized (verbatim based processing). Indeed, this aspect proved to be true for a sample of monolingual adults. Little is known about what happens for bilinguals, for whom emotional word processing may vary on the basis of one's linguistic knowledge.

Therefore, we intended to verify whether this verbal knowledge is able to affect semantic processing or whether, on the contrary, this competence develops so deeply and detached from language as to allow a profound elaboration of the verbal material, influenced by the emotionality of the terms, like for monolingual participants.

Our first hypothesis was that bilinguals might have a pattern of false memories similar to that of their monolingual peers. In fact, we thought that their semantic processing might be similar to the one of the monolingual control group, causing a similar degree of distortion.

As a second hypothesis we expected fewer / smaller differences caused by the emotion conveyed by the critical lures. In particular, on the basis of previous results we found on other bilinguals (Cangelosi, Bossi & Palladino, 2021) we expected less differences concerning the valence of the critical lures, since valence seems to be particularly sensitive to the specificities of each language.

Method

Participants.

The sample was made up of 44 participants (Average age: 27.2, Range: 22-35), 19 of which were bilinguals with German as L1 and Italian as L2 and 25 were monolinguals Italian. Both groups signed the informed consent to take part in the experiment and filled out the Four Factor Index of Social Status Questionnaire (Hollingshead, 1975), in a version adapted to adults. Furthermore, the bilingual group completed a questionnaire aimed at assessing their linguistic background (based on Barac & Bialystok, 2012, on Luk

& Bialystok, 2013 and on Anderson, Mak, Chahi and Bialystok, 2018), considering their mother tongue, their language spoken at home, their age of acquisition of the L2, their language of main exposure and their context of learning and use of their two languages spoken. More in detail, after some general information regarding the registry part, age of first exposure was assessed through an open question and through three further multiple choice questions; the frequency of use of the two languages was evaluated through eight questions in which participants had to evaluate some common situations and the proportion of use in each of them on a Likert scale from 1 (= never) to 10 (= always).

Materials.

All the materials have been presented and compiled via computer.

Cornell/Cortland Emotional DRM Lists (Brainerd et al., 2008).

In order to measure verbal false memory, our participants were exposed to 12 emotional DRM lists, presented in Italian, L2 for the German bilinguals and L1 for the monolinguals.

The materials have been adapted and validated[ii] from the original Cornell/Cortland Emotional lists (CEL Lists, 2008), following the same testing procedure used by the Authors on a sample with comparable age. In order to adapt and validate the Cornell/Cortland Emotional Lists (CEL Lists, 2008) to Italian we've conducted a pilot study that involved 20 young adults (Mean age: 26.5), who did not take part in the real study. The frequencies of use of the words were considered referring to the ColFis Lexical Database (Bertinetto *et al*., 2005). The Backward Associative Strength index (BAS, i.e., how much each word included in the list is associated with the critical lure)

was evaluated on a 7-point Likert scale (from 1 = not at all associated to 7 = very associated).

The words have been ordered on the basis of the assigned BAS that we derived from our pilot study on the Italian lexicon, from the most to the less associated.

In accordance with the procedure followed by the authors, twelve lists have been chosen from the adapted-to-Italian 32 Cornell/Cortland Emotional Lists (CEL Lists, 2008). Every list presented in the coding phase was made up of 10 of the 15 original words, ordered from the one with the stronger BAS, to that with the weaker BAS obtained through our pilot study. The 5 words with the lowest BAS value, out of the 15 words of each list, were thus eliminated. The 12 lists were 3 from each of the four Arousal * Valence combination: High Arousal * Negative Valence, Low Arousal * Negative Valence, High Arousal * Positive Valence, Low Arousal * Positive Valence.

The recognition task was related to every Arousal * Valence combination and was made up of 21 words: 9 target words (3 from each presented list), the 3 not presented critical lures corresponding to each presented list, 1 related distractor for each presented list (taken by the 12th position of the 15-words lists) and 6 unrelated distractors, two for each of three unpresented lists, belonging to the same Arousal * Valence combination.

Vocabulary

In order to assess receptive vocabulary skills, we proposed a verbal meaning task (PMA, 1981) in which participants had to correctly mark the synonym of a given word, choosing between four alternatives. We proposed this task in a short version developed by Cavallini and colleagues (*in preparation*) , in which participants had 3 minutes to complete the task (instead of the 8 minutes expected in the original

version), composed by the 18 more representative trials of the original 30s; scoring consists in counting the number of correct responses.

Categorical Fluency

We administered a fluency task to assess productive vocabulary: participants had to name as many words as possible belonging to a given category (i.e., “car brands”, “fruits” and “animals”); they had one minute for each of the three categories. Scores were obtained considering the total correct words produced for each category in the given time, as in Krumm and colleagues (2021).

Forward/Backward Digit-Span Task (Soylu, 2010)

We administered the two versions of this task in order to measure short-term memory, here represented by numeric span. Every participant was verbally presented with series of numbers, divided into blocks of 3 sequences each. Within each block, the sequences had the same length; the blocks were divided according to the gradually increasing length of the sequences they contain.

In the first of the two tasks (i.e. Forward Digit-Span Task) the participant was instructed to repeat each sequence, after listening to it, in the same order in which it was presented (e.g., the experimenter reads: "1-3-5"; The participant repeats: "1-3-5").
On the contrary, in the second of the two tasks the instruction was to repeat each sequence, at the end of the oral presentation of the same, in the reverse order compared to that according to which it was read (e.g., the experimenter reads: "5-7-9" ; the participant replies: "9-7-5").

In the version of the two tasks administered here, the task ends when the participant missed two out of three sequences within a block. The quantity of numbers of

the last correctly executed block constitutes the score obtained in the task. For example, a score of 5 will be obtained for the Forward part, if the participant misses two sequences in the 6-digit block; similarly, for example you will have a score of 4 obtained for the Backward part, if the participant misses two sequences in the 5-digit block.

Procedure

Participants took part in a telematic meeting, through the Skype platform, lasting about one hour.

At first they had to sign the informed consent via Google form, through a link that the experimenter shared by chat. The informed consent was mandatory to go on with the meeting. Below the informed consent there were also the two questionnaire sections, the one regarding the socioeconomic status and the one regarding the linguistic use and background. Participants were asked to share their screen, in order to read the consent and the other two questionnaires together with the experimenter, allowed to guide them or to answer step by step possible questions. The administration of our control measures was randomized between participants, in order to avoid the effect of fatigue.

Cornell/Cortland Emotional DRM Lists (Brainerd et al., 2009).

The experimental task, built through OpenSesame 2.0 program (Mathôt, Schreij, & Theeuwes, 2012), was proposed through the Just Another Tool for Online Studies platform (JATOS, Lange, Kuhn & Filevich, 2015), which allowed to share the task with the participants via link. Participants had to share their screen, and the instructions appeared contemporary on their screen, being integrated by the experimenter. A shirt practice containing a neutral DRM lists to be decoded was run before the real experiment started.

The procedure adopted in the administration was the same proposed by the authors of the original lists (Brainerd et al., 2008). During the coding phase, three lists appeared on the screen, one word at a time, with a 10 second gap between one list and another. The stimuli were written in “mono” font, in the center of the screen. Participants were asked to focus on the words and to memorize as many of them as possible because they would later on be asked to recognize them among other not coded words. After the first three lists there was the first recognition task was. Here participants had to press the space bar every time they thought to recognized one of the coded words, as fast and as accurately as possible. Between the coding phase and the recognition task there was a 30-seconds pause, in which participants were asked some generic questions in order to distract them from the words. The coding-and-recognition procedure was repeated for four times, each with different lists. The lists composing each block all belonged to one of the four valence X arousal conditions. The order of the valence X arousal conditions was randomized between participants, so as for the order of the lists in each block and so as of the words composing every recognition phase.

Vocabulary

The vocabulary task was administered as a Google Form link, in which participants had to select their answer between four multiple choices. As in the original paper format version, they had the chance to skip one item if they were not sure about the answer and fill it later, or leave it blank. As for the questionnaire, they had to share their screen while completing the task. Before the experimenter stared the stopwatch, the instructions were read aloud together with the participant and they were accompanied by an example.

Categorical Fluency

Concerning the categorical fluency task, the experimenter gave aloud the instructions to every participant, providing an example of words belonging to a category that was not included in the real task. Participants, then, had to name aloud as many words as possible that they thought could fit with the given category, in one minute of time. The experimenter directly wrote the words on a word document. This procedure was repeated for the three categories, with shirt breaks between one and the other.

Forward/Backward Digit-Span Task (Soylu, 2010)

Concerning both the Forward and the Backward version of the Digit-Span Task, the experimenter read the sequences of numbers aloud, writing down the participant's answers. The sequences were read at a rate of approximately 1 second one before the other. Both in the case of the Forward Digit-Span and in the case of the Backward Digit-Span, a trial was provided as example before starting the testing.

Results

Data planning

All analyses were done using the package "lme4" in R Studio 1.0.153 (RStudioTeam, 2020) for mixed effect models.

As a first analysis, we ran a mixed model on a global sample, on the basis of the nature of our dependent variable, consisting on a dichotomous "yes/no" response, which allows us to analyze a binomial response. We considered both monolinguals and bilinguals, with d' as our dependent variable, with bilingualism in interaction with valence and arousal of the critical lures as main effects, with vocabulary, fluency, span

and reaction times (RTs) as covariates and, to conclude, with valence + arousal per participant as our random effects.

Analysis

From this first analysis it emerged that valence, the interaction between valence and arousal, fluency and the RTs were significant ($p = 0.005$, $\chi2$ (1) = .347; $p = 0.011$, $\chi2$ (1) = 6.293, $p = .036$, $\chi2$ (1) = 4.670, $p = .030$, $\chi2$ (1) = 4.848), while arousal had a trend towards significance ($p = 0.085$, $\chi2$ (1) = 1.286). The span, considered here as an aggregate index of direct and backward span, was not significant ($p = .825$, $\chi2$ (1) = .127). Bilingualism, in turn, was not significant either in itself ($p = .870$, $\chi2$ (1) = .315), or in relation to the emotional variables of interest (Bilingualism * Arousal, $p = .230$, $\chi2$ (1) = 1.985; Bilingualism * Valence, $p = .960$, $\chi2$ (1) = .167; Bilingualism * Valence * Arousal, $p = .500$, $\chi2$ (1) = .516) (See Figure 3.1).

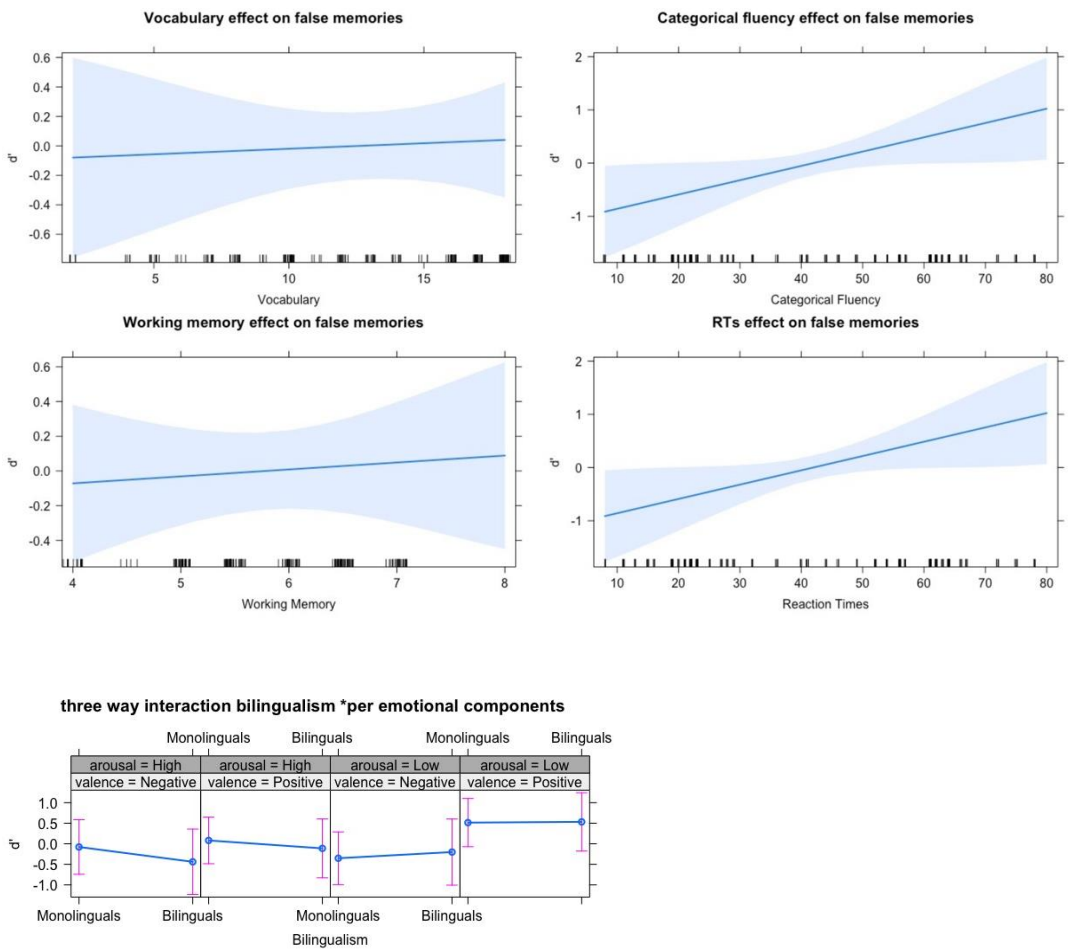


Figure 3.1. Results of the Mixed Model of the general sample.

The analysis thus confirmed that bilingualism is not predictive of a different performance in terms of false memory, neither in an absolute sense and isolated from emotional variables, nor in single or double interaction with them. As a result, it seemed that the semantic processing of the bilinguals making up this sample develops in a way that is comparable to what happens for monolingual peers, both considering it in itself and in its connotative dimension. This result should be considered net of a significantly worse average performance in terms of vocabulary for the bilingual group ($p < .01$).

For the sample as a whole, valence and its interaction with arousal played a primary role (see Figures 3.2-3.4).

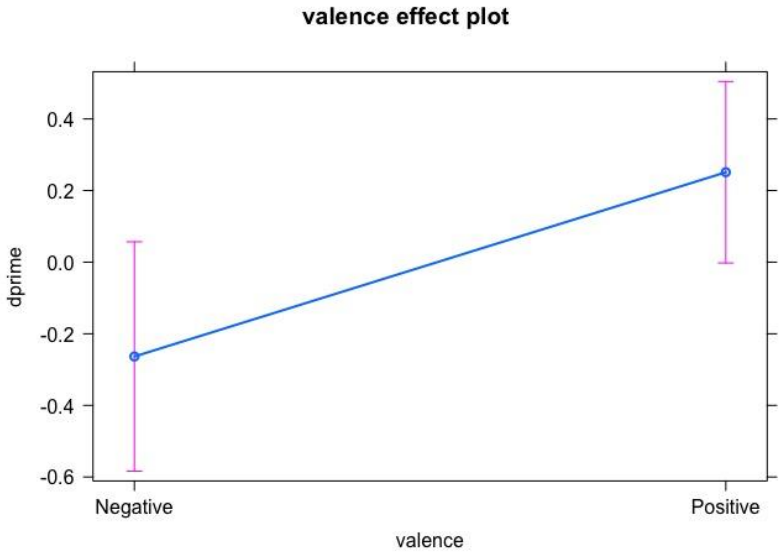


Figure 3.2. Valence effect on d' results. The critical lure with negative valence elicit the higher number of false memories (that correspond to lower d' values); on the contrary, the critical lure with positive valence elicit a lower number of false memories in the general sample composed by both monolinguals and bilinguals.

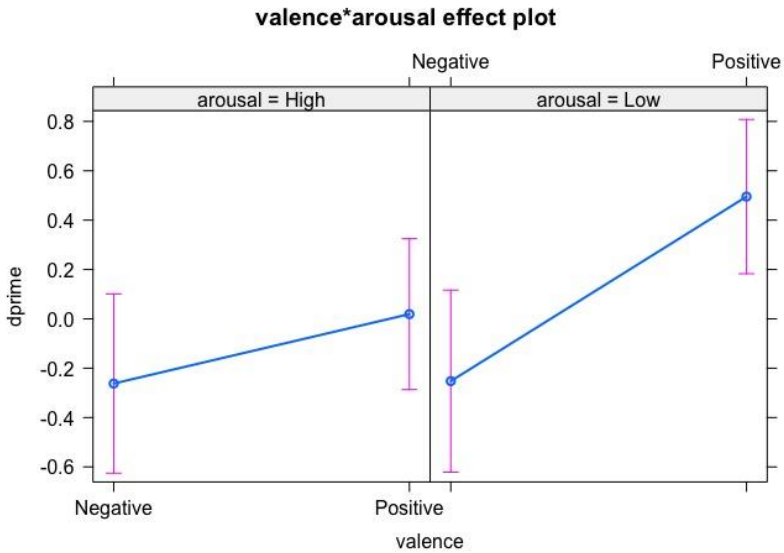


Figure 3.3. Here it is possible to observe how the arousal of the critical lures affects d' values, according with the two possible valences (negative and positive). Both in a condition of high and low arousal of the critical lures it is possible to observe the pattern according to which negative valence affects the number of false memories more than the positive valence does. Though, it is possible to highlight that in the high arousal condition there Is less difference between the effect of negative and positive valence of the critical lures. Indeed, in this arousal condition there is a higher number of false memories as a whole with respect to the low arousal condition in which the words with negative valence are more distant from the positive ones as regards the probability to elicit false memories.

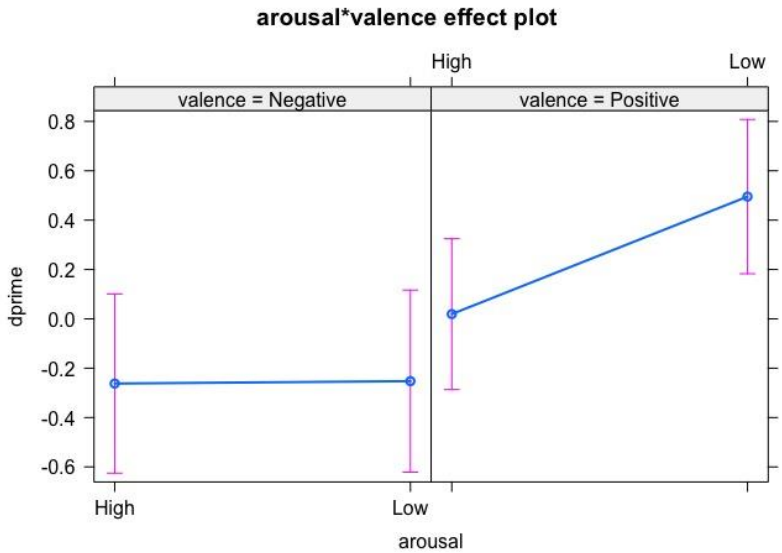


Figure 3.4. Here it is possible to observe the reverse of the effect described in Figure 3, that is how the valence of the words changes according to their arousal. In particular it is possible to observe that there seem not to be a difference between high and low arousal when considering negative valence of the words. On the contrary, when considering the positive valence of the words, arousal is able to affect the performance in our recognition task. In particular, critical lures with positive valence are more often recognized as already seen in a high arousal condition then in a low arousal condition.

In particular, negative valence lists are more capable of distorting memory and eliciting false memory for semantically associated lures.

As for the arousal, it can be seen how, in accordance with the literature, highly emotional lists cause a greater number of false memories than less emotionally activating lists, even if this effect is not significant in itself but only in interaction with valence. As a further step and in light of the significance of the RTs in our general model, we decided to investigate their effect, carrying out the analysis separately on the actually coded words and on the critical lures. We decided again to use a mixed-effect model approach.

With respect to the model run only on the actually coded words, the number of words that were correctly recognized was our dependent variable. Bilingualism was inserted in interaction with valence and arousal, all included in the model as main effects.

Vocabulary, categorial fluency and numeric span were inserted as covariates. Valence + Arousal per participant were inserted as random effects.

Being bilingual was significant in this model ($p < .01$, $\chi 2$ (1) = 14.683) with a negative estimate, meaning that being bilingual in our sample leads to less accuracy. The RTs were significant ($p < .01$, $\chi 2$ (1) = 228.254) with a negative estimate, meaning that a shorter reaction time in this case corresponds to a greater accuracy. Valence and arousal were not significant in this model: this aspect is consistent with the fact that only the lures in our word pool are emotionally connoted according to the two axes valence and arousal (the words actually presented are linked to the lures by semantic area, coded and represented by the BAS index, backward association strength) (See Figure 3.5).

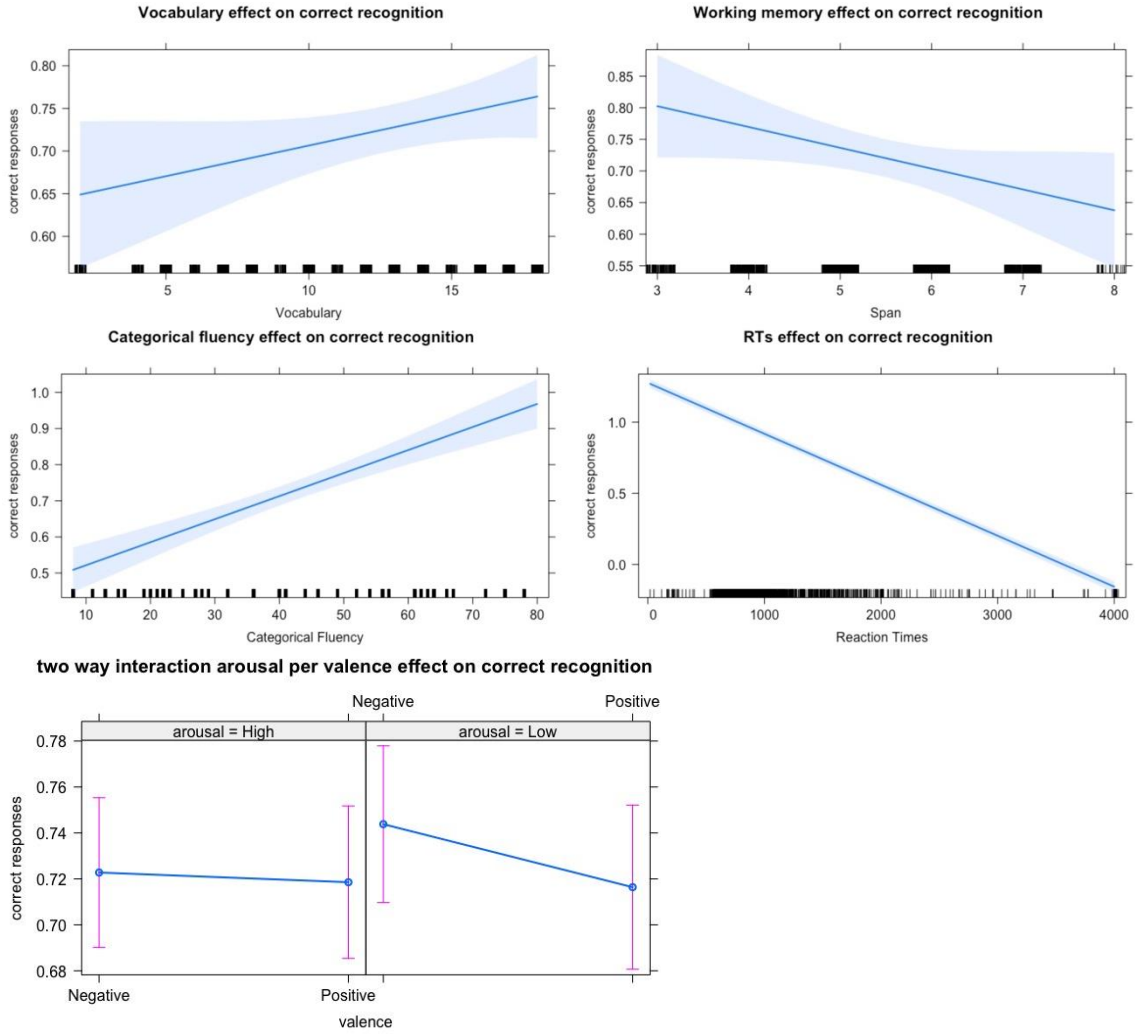


Figure 3.5. Results of the Mixed Model run exclusively on the words that were actually present during the coding phase.

Similarly, concerning the model run only on the semantically related critical lures we inserted the number of lures recognized as already been coded even if they were not as dependent variable. Bilingualism in interaction with valence and arousal were our main effects. Vocabulary, Fluency and Span were our covariates, while valence + arousal per participant were our random effects.

Bilingualism was significant in this model ($p < .01$, $\chi2$ (1) = 39.488) with a negative estimate, meaning that being bilingual in our sample leads to a reduction in terms of semantically related false memories for the critical lures. The RTs were significant ($p < .01$, $\chi2$ (1) = 158.348) with a negative estimate, meaning that a shorter

reaction time in this case corresponds to a greater production of falsely remembered lures. Positive valence was significant ($p < .01$, $\chi 2$ (1) = 3.044) with a negative estimate meaning that words with a positive valence lead to a lower production of false memories. Similarly, the low arousal of the words was significant ($p < .01$, $\chi 2$ (1) = 3.014) with a negative estimate meaning that the less emotionally connoted words lead to a lower production of false memories. Both of these aspects were consistent with the literature that affirms how the positive valence and low emotionality of words are associated with greater accuracy and verbatim based processing rather than on gist based characteristics, more semantically linked to the core of the information conveyed by the stimulus.

Furthermore, the interaction between bilingualism * positive valence * low arousal proved to be significant ($p < .01$, $\chi 2$ (1) = 1.277) with a negative estimate, meaning that bilinguals in particular are less influenced by the semantic activation elicited by the lures when they have to encode words with positive valence and low arousal (See Figure 3.6).

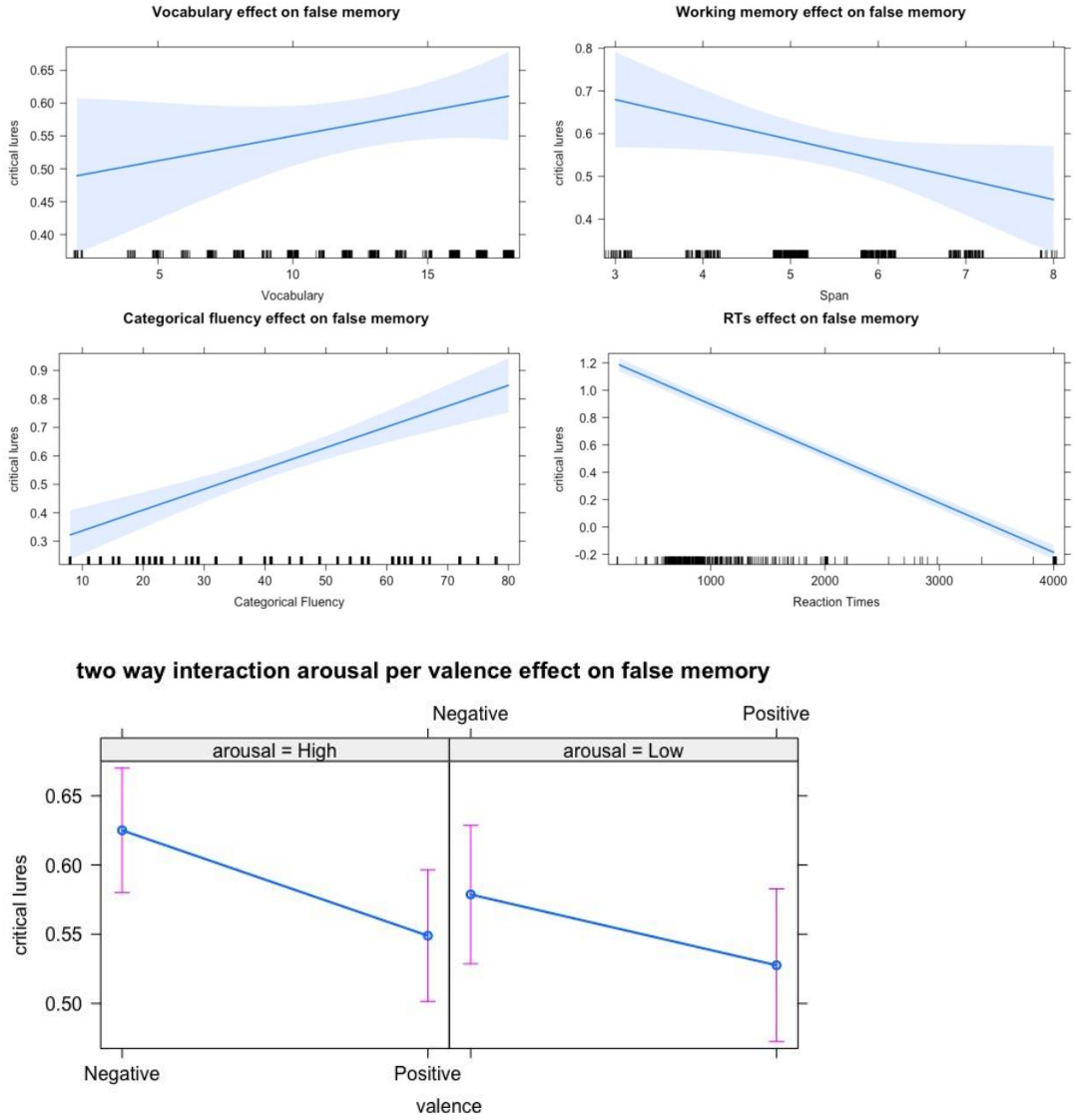


Figure 3.6. Results of the Mixed Model run exclusively on the semantically related critical lures that were not present during the coding phase.

Discussion

We reviewed the literature concerning the effect of emotionality on verbal false memories', and we found a gap concerning how it acts in bilinguals' L2. We decided to explore how the emotional content of a pool of DRM lists influences verbal false memory

in the second language of a sample of German/Italian bilingual adults, in comparison with a group of monolinguals peers, similar in terms of education degree.

It has been demonstrated on monolingual groups that the greater emotionality of the actually coded terms enhances semantic based (or gist based) processing, to the detriment of a more precise analysis of the specific characteristics of the material that is required to be memorized (verbatim based processing). Our aim was to verify if the words in a bilingual's L2 can convey emotionality as they do for monolinguals, so as to influence memory processing and semantic elaboration.

Our first hypothesis was that bilinguals might have been comparable to monolinguals in terms of number of false memories. Indeed, we thought that they might be affected by semantic processing similarly to monolinguals.

Concerning this first hypothesis, from the analysis run on the actually coded words, it turned out that bilinguals were less accurate than monolinguals. This result is particularly interesting since in opposition they also showed a significant reduction on verbal false memories in the analysis run only on critical lures.

As a second hypothesis we expected that emotionality of the items might be less prominent in causing verbal false memories in bilinguals', in light of a possible reduced emotionality in their L2.

With respect to this second hypothesis, it is possible to state that there seems not to be differences with respect to the role of arousal in causing verbal false memories in itself between the two groups. It has to be noticed that arousal turned out not to be significant in the general analysis in itself, but only in relation with valence. However, its interaction with being bilingual was not significant, meaning that its role can be meant as equivalent in these two groups. In particular and coherently with the literature (e.g., Brainerd *et al*.,

2010) higher arousal lures are more prone to be falsely recognized as already seen than lower arousal ones in both the groups.

This role of the arousal was confirmed by its significance in the analysis run only on critical lures, indicating that lower arousal lures were effective in reducing verbal false memories both for bilinguals and for monolinguals.

Valence was significant in the analysis run on all the items, in itself and in interaction with arousal, but regardless of being bilingual. In particular, coherently with the literature (e.g., Brainerd *et al.*, 2010) negative valence items were significant in causing an higher number of verbal false memories both for bilinguals and for monolinguals. This effect was particularly strong for the higher arousal and negative valence lures.

Concerning the role of valence in the analysis run on the critical lures, positive valence lures were less falsely remembered than negative valence ones, consistently with the findings of the analysis run on all the items. It also emerged that bilinguals are particularly sensitive to the combination of positive valence and low arousal lures, less falsely recognized especially in this group.

With respect to the control measures, categorical fluency turned out to be significant in affecting verbal false memories. In particular, persons with higher scores in categorical fluency produced an higher number of verbal false memories.

RTs were also significant. In particular, lower RTs corresponded to a higher number of verbal false memories.

These two results seem to demonstrate ones more something regarding the mechanisms underlying this kind of verbal false memories. In fact, the role of categorical fluency demonstrate that verbal false memories arise in light of processing the information on the basis of the meaning that it can convey. On the other hand, RTs point out their very

nature of distortion of memory to the detriment of accuracy. In sums, these two aspects confirm the role of verbal false memory as a sign of memory fallibility for those punctual features that determine accuracy but able to identify a particular way of semantically processing the information.

We can compare these results to the ones obtained in our previous study (Cangelosi, Bossi, & Palladino, 2021) on minority language bilingual children. Even if the design of the two studies are slightly different in terms of covariates and especially in terms of age of the groups involved, nevertheless, it is possible to make some general inferences.

In fact, these two bilingual populations especially differ in terms of SES and status of their first language, as recognized by their linguistic background. The results of these two studies are very different in that the first one run on minority language bilingual children allowed us to state that there were no differences in semantic processing and in the role conveyed by emotionality of the words to be memorized. On the contrary, from the present study we can argue that being bilingual as German/Italian bilinguals of Bozen are leads to differences both at the level of semantic processing in itself, and at the level of connotation, namely the relevance that emotionality has on that semantic processing. In sums, these results tells that it is not possible to make a general inference on being bilingual as an absolute concept, but that there are some determinant and huge differences in the way in which one can be considered as bilinguals. These differences lead to differences in cognitive and linguistic processes.

As a future study we would like to run this experimental design on a sample composed by German/Italian bilingual living in Bozen of the same age of our 2021's study, namely we would like to recruit children aged from 9 to 11 years so as to have a stronger comparison to be made between these two populations. In particular, we expect children to be more similar to the previous one in terms of emotional role, since we

expect them to be nearer to simultaneous bilingualism at that age than as adults. In fact, as adults they tend to choose a dominant language, that in their case is German and to lose maybe further connections with their L2. On the contrary, minority language bilingual children of our 2021 study are expected to have an opposite trajectory, since living in Italy they will have to deepened their Italian language knowledge more and more, maybe at the expense of their mother tongues. Indeed, their second language is expected to become increasingly emotionally connotated, while the German/Italian bilinguals may lose the emotionality that Italian is able to convey in their case, due to practice and daily use it less as an average.

To briefly sum up the main results of the present study, we addressed the role of emotionality of the words (intended as their arousal and their valence) has on semantic processing that is at the basis of verbal false memories comparing a group of German/Italian bilinguals from Bozen with a group of Italian monolinguals. We found that these bilinguals have a reduced number of verbal false memories and also a reduced accuracy in true memory, meaning that reduction is relevant in terms of semantic processing in their L2. Furthermore, they produced significantly less false memories for the combination of low arousal and positive valence of the critical lure. Although this result is coherent with the previous literature on monolinguals (e.g., Brainerd *et al.*, 2010), this emotional combination seems to be particularly effective in avoiding their verbal false memories.

References

Altarriba J and Basnight-Brown DM (2009) An overview of semantic processing in bilinguals: Methods and findings. In Pavlenko A (ed.), The bilingual mental lexicon: Interdisciplinary approaches. *Clevedon: Multilingual Matters*, pp. 79–99. DOI: 10.22158/eltls.v2n3p43

Anderson, J. A., Mak, L., Chahi, A. K., & Bialystok, E. (2018). The language and social background questionnaire: Assessing degree of bilingualism in a diverse population. *Behavior research methods, 50*(1), 250-263.

Anooshian, L. J., & Hertel, P. T. (1994). Emotionality in free recall: Language specificity in bilingual memory. *Cognition and Emotion, 8*(6), 503–514. https://doi.org/10.1080/02699939408408956

Ayçiçeği, A., & Harris, C. L. (2004). Brief Report: Bilinguals' recall and recognition of emotion words. *Cognition and Emotion, 18*(7), 977-987. https://doi.org/10.1080/02699930341000301

Barac, R., & Bialystok, E. (2012). Bilingual effects on cognitive and linguistic development: Role of language, cultural background, and education. *Child development, 83*(2), 413-422.

Bertinetto, P. M., Burani, C., Laudanna, A., Marconi, L., Ratti, D., Rolando, C., & Thornton A. M. (2005). Corpus e Lessico di Frequenza dell'Italiano Scritto (CoLFIS*)*.

Brainerd, C., & Reyna, V. (2004). Fuzzy-trace theory and memory development. *Developmental Review, 24*(4), 396–439. https://doi.org/10.1016/j.dr.2004.08.005

Brainerd, C. J., Yang, Y., Toglia, M. P., Reyna, V. F., & Stahl, C. (2008). Emotion and false memory: The Cornell/Cortland norms. In *annual meeting of the Psychonomic Society, Chicago, IL.*

Brainerd, C. J., Yang, Y., Reyna, V. F., Howe, M. L., & Mills, B. A. (2008). Semantic processing in "associative" false memory. Psychonomic Bulletin & Review, 15(6), 1035–1053. https://doi.org/10.3758/pbr.15.6.1035

Brainerd, C., Stein, L., Silveira, R., Rohenkohl, G., & Reyna, V. (2008). How Does Negative Emotion Cause False Memories? *Psychological Science, 19*(9), 919–925. https://doi.org/10.1111/j.1467-9280.2008.02177.x

Brainerd, C., Holliday, R., Reyna, V., Yang, Y., & Toglia, M. (2010). Developmental reversals in false memory: Effects of emotional valence and arousal. *Journal of Experimental Child Psychology, 107*(2), 137–154. https://doi.org/10.1016/j.jecp.2010.04.013

Braunstein, V., Ischebeck, A., Brunner, C., Grabner, R.H., Stamenov, M. & Neuper, C. (2012) Investigating the influence of proficiency on semantic processing in bilinguals: An ERP and ERD/S analysis. *Acta neurobiologiae experimentalis 72*, 421–38.

Cangelosi, M., Bossi, F., & Palladino, P. (2021). Did you see that? False memories for emotional words in bilingual children. *Bilingualism: Language and Cognition*, 1–13. https://doi.org/10.1017/s136672892100105x

Deese, J. (1959) On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology 58*, 17–22. DOI: 10.1037/h0046671

Dehon, H., Larøi, F., & van der Linden, M. (2010). Affective valence influences participant's susceptibility to false memories and illusory recollection. *Emotion, 10*(5), 627–639. https://doi.org/10.1037/a0019595

Dong, Y., Gui, S., & MacWhinney, B. (2005). Shared and separate meanings in the bilingual mental lexicon. Bilingualism: Language and Cognition, 8(3), 221–238. https://doi.org/10.1017/s1366728905002270

El Sharkawy, J., Groth, K., Vetter, C., Beraldi, A., & Fast, K. (2008). False memories of emotional and neutral words. *Behavioural Neurology, 19*(1-2), 7–11. https://doi.org/10.1155/2008/587239

Francis, W.S. (2018) Shared core meanings and shared associations in bilingual semantic memory: Evidence from research on implicit memory. *International Journal of Bilingualism 24*, 464-477. DOI: 10.1177/1367006918814375

van Heuven, W. J. B., Dijkstra, T., & Grainger, J. (1998). Orthographic neighborhood effects in bilingual word recognition. *Journal of Memory and Language, 39*(3), 458–483. https://doi.org/10.1006/jmla.1998.2584

Howe, M. L., & Derbish, M. H. (2010). On the susceptibility of adaptive memory to false memory illusions. *Cognition, 115*(2), 252–267. https://doi.org/10.1016/j.cognition.2009.12.016

Howe, M. L., Wimmer, M. C., Gagnon, N., & Plumpton, S. (2009). An associative-activation theory of children's and adults' memory illusions. *Journal of Memory and Language*, *60*(2), 229–251. https://doi.org/10.1016/j.jml.2008.10.002

Khateb, A., Pegna, A.J., Michel, C.M., Mouthon, M. & Annoni, J.M. (2016) Semantic relatedness and first-second language effects in the bilingual brain: A brain mapping study. *Bilingualism: Language and Cognition 19*, 311–330. DOI: 10.1017/S1366728915000140

Krumm, S., Berres, M., Kivisaari, S. L., Monsch, A. U., Reinhardt, J., Blatow, M., Kressig, R. W., & Taylor, K. I. (2020). Cats and Apples: Semantic Fluency Performance for Living Things Identifies Patients with Very Early Alzheimer's Disease. *Archives of Clinical Neuropsychology*. Published. https://doi.org/10.1093/arclin/acaa109

Lange K, Kühn S, Filevich E (2015) Correction: "Just Another Tool for Online Studies" (JATOS): An Easy Solution for Setup and Management of Web Servers Supporting Online Studies. PLOS ONE 10(7): e0134073. https://doi.org/10.1371/journal.pone.0134073

Luk, G., & Bialystok, E. (2013). Bilingualism is not a categorical variable: Interaction between language proficiency and usage. *Journal of Cognitive Psychology, 25*(5), 605-621.

Mathôt, S., Schreij, D., & Theeuwes, J. (2011). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods, 44*, 314–324. DOI: 10.3758/s13428-011-0168-7

Novitskiy, N., Myachykov, A. & Shtyrov, Y. (2018) Crosslinguistic interplay between semantics and phonology in late bilinguals: Neurophysiological evidence. *Bilingualism: Language and Cognition 22*, 209–227. DOI: 10.1017/S1366728918000627

Palmer, J. E., & Dodson, C. S. (2009). Investigating the mechanisms fuelling reduced false recall of emotional material. *Cognition & Emotion, 23*(2), 238–259. https://doi.org/10.1080/02699930801976663

Reyna, V.F., Corbin, J.C., Weldon, R.B. & Brainerd, C.J. (2016) How fuzzy-trace theory predicts true and false memories for words, sentences, and narratives. *Journal of applied research in memory and cognition 5*, 1–9.

Roediger, H. L., Balota, D. A., & Watson, J. M. (2001). Spreading activation and arousal of false memories. In R.G. Crowder and H.L. Roediger (Eds.), *The nature of remembering: Essays in honor of Robert G. Crowder* (pp. 95-115). Washington DC: American Psychological Association.

Roediger, H.L. & McDermott, K.B. (1995) Creating false memories: Remembering words not presented in list. *Journal of Experimental Psychology: Learning, Memory, and Cognition 21,* 803–814. DOI: 10.1037/ 0278-7393.21.4.803

RStudio Team (2020). *RStudio: Integrated Development for R. RStudio, PBC, Boston, MA URL http://www.rstudio.com/*

Soylu, F. (2010). Forward/backward digit-span task. *Arch. Neurobehav. Exp. Stimuli* 222. http://www.neurobs.com/ex_files/expt_view?id=218

Zhang, W., Gross, J., & Hayne, H. (2016) The effect of mood on false memory for emotional DRM word lists. Cognition and Emotion 31, 526–537. DOI: 10.1080/02699931.2016.1138930

Section 2

Verbal false memories in older adults

Theoretical issue

In this second section we will focus on Verbal False Memory in older adults, a population with characteristics that mirror that of bilinguals. Here we will discuss what is going to happen in chapter 5 and 6. We will set the stages by discussing the literature background and anticipating the materials and experimental designs. When bilinguals are compared to monolinguals, they possess enhanced executive functions (including working memory) and a reduced vocabulary in each of the two languages, at least in the first phase of foreign language learning (Bialystok, 2009). In contrast, bilinguals a pattern that is opposite to that of older adults. Executive functioning and memory are both affected by aging and tend to decline over the years. Vocabulary size, however, tends to remain stable and even to increase according to some studies (e.g., Birren, 2006).

However, also in this population, there are really few evidence regarding the effect of the emotional-semantic component of the lexicon on the production of false verbal memories. In fact, we found only one study (i.e., Kensinger & Corkin, 2004) addressing the role of emotionality on verbal false memory in aging and this study didn't consider the role of valence but only the one of arousal.

Before investigating how the connotative component of semantics acts on the production of false memories, we must consider a further gap upstream. Older adults have been found to produce more false memories than other age groups (e.g., Devitt & Schacter, 2016), but it is not completely clear if this is linked to semantics, as for the other age groups (e.g., Brainerd, 2008), or if it is mainly due to a decrease in executive functions for this age group.

Verbal false memories in aging: two theoretical views

Some researchers argue that the increased production of false verbal memories in aging arises because older adults tend to process information semantically, mainly processing the gist of what they are asked to code rather than relying on verbatim characteristics. (e.g., Tun et al., 1998; Dennis, Kim, and Cabeza, 2007). From the available studies it is inferred that the older adults use semantic processing to compensate for the decline in other cognitive domains (i.e., memory and executive functioning).

Other researchers, however, argue that the increased production of false memories in older adults is a consequence of inhibitory processes in relation to executive functioning. For example, Askey and Playfoot (2018) explicitly criticized the view of enhanced semantic processing in aging, saying that if it should be so there should be not only an enhancement in false memory but also in true memory. On the contrary, on the basis of their results they suggested that the Inhibition Deficit Hypothesis can account for the increase in verbal false memory in aging. This theory states that the increase in false memory might be due to the mechanism in aging becoming less efficient.

Participant profile

We compared young participants between 20 and 35 years old with older adults from 65 to 80 years, including both in our experimental design. The group of older participants was administered the Mini Mental State Examination (MMSE, Folstein & McHugh, 1975), which assesses a person's neuro-cognitive and functional state. Only participants passing the 24-point *cutoff*, usually used as a benchmark for distinguishing healthy aging from pathological aging, were included in the older adults group.

Task design

In this study we included and compare the two above mentioned age groups, with the aim to verify the contribution of semantics as an independent factor, before making any inferences with respect to emotional aspects. To do this, an index of semantic similarity was considered, which will be discussed in more detail in the course of the study. As for the studies presented in the previous section, we also used the Cornell/Cortland Emotion Lists (CEL Lists; Brainerd et al., 2008) here, but considering the lists, however, devoid of the emotional aspect.

Internal DRM measure

Semantic Similarity (SSim)

To test the effect of semantics on verbal false memories, we computed a semantic similarity index for the critical lures and distractors present in the DRM paradigm. Vector representations for these words were extracted from the semantic space WEISS1 – Italian forms based on the Continuous Bag of Words (CBOW) model (Mikolov, Chen et al., 2013), trained on the Italian-text corpus ItWac.

Additional measures

In this case, we will talk about additional measures, rather than about control variables. The goal of including these measures was to provide information on the basis of the production of false verbal memories in the older adults, adding to information furnished by the DRM paradigm itself.

As extra external tasks, we included Short-Term memory / working memory (Direct / Reverse Span) and measured inhibition skills (Hayling Task) as an aspect of executive functioning.

The role of emotionality on verbal false memories in older adults

Although the emotional component was not considered for the purposes of the main study and needs further investigation, for completeness we chose to include a short section of additional analyses relating to the role of valence and arousal on verbal false memories in aging. We can consider these two dimensions for different age groups, because the Cornell / Cortland emotional lists were also used in the previous study. For the purpose of this additional study, we will also consider the CES-D questionnaire for measuring depression, a vocabulary measure, and a categorical fluency measure (included in the analysis as an aggregate vocabulary index). This was combined with Direct and Backward Span tasks (included in the analysis as an aggregate index of Span) and the Hayling Task to check for executive functioning.

References

Askey, C., & Playfoot, D. (2018). Examining theories of cognitive ageing using the false memory paradigm. *Quarterly Journal of Experimental Psychology, 71* (4), 931–939. https://doi.org/10.1080/17470218.2017.1307433

Bialystok, E. (2009). Bilingualism: The good, the bad, and the indifferent. *Bilingualism: Language and Cognition, 12* (1), 3–11. https://doi.org/10.1017/s1366728908003477

Birren, J. E. *Handbook of the Psychology of Aging, Sixth Edition (Handbooks of Aging) (2005–12-23).* (2021). Academic Press; 6 edition (2005–12-23).

Brainerd, C. J., Yang, Y., Reyna, V. F., Howe, M. L., & Mills, B. A. (2008). Semantic processing in "associative" false memory. *Psychonomic Bulletin & Review, 15* (6), 1035–1053. https://doi.org/10.3758/pbr.15.6.1035

Brainerd, C. J., Yang, Y., Toglia, M. P., Reyna, V. F., & Stahl, C. (2008). Emotion and false memory: The Cornell/Cortland lists. In *annual meeting of the Psychonomic Society, Chicago, IL*.

Dennis, N. A., Kim, H., & Cabeza, R. (2007). Effects of aging on true and false memory formation: An fMRI study. *Neuropsychologia, 45* (14), 3157–3166. https://doi.org/10.1016/j.neuropsychologia.2007.07.003

Devitt, A. L., Tippett, L., Schacter, D. L., & Addis, D. R. (2016). Autobiographical memory conjunction errors in younger and older adults: Evidence for a role of inhibitory ability. *Psychology and Aging, 31*(8), 927–942. https://doi.org/10.1037/pag0000129

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state." *Journal of Psychiatric Research, 12* (3), 189–198. https://doi.org/10.1016/0022-3956(75)90026-6

Kensinger, E. A., & Corkin, S. (2004). The effects of emotional content and aging on false memories. *Cognitive, Affective, & Behavioral Neuroscience*, *4*(1), 1–9. https://doi.org/10.3758/cabn.4.1.1

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. Retrieved from https://arxiv.org/abs/1301.3781.

Tun, P. A., Wingfield, A., Rosen, M. J., & Blanchard, L. (1998). Response latencies for false memories: Gist-based processes in normal aging. *Psychology and Aging, 13* (2), 230–241. https://doi.org/10.1037/0882-7974.13.2.230

CHAPTER 5

Older adults produce more verbal false memories than younger adults: is it semantics or executive functioning?

In preparation: Cangelosi, M., Rinaldi, L., Dijkstra, T., Palladino, P., & Cavallini, E. (2022).
To be submitted to *Memory and Cognition.*

Abstract.

A verbal false memory occurs when one 'remembers' a word to be part of a previously presented list of words, while this was not in fact the case. It may occur as a consequence of the semantic associations of the word with words actually part of the list. False memories for people of all ages have been extensively studied by means of the Deese-Roediger-McDermott paradigm (DRM paradigm). Results show older adults suffer more from verbal false memories than other age groups. One possible explanation of this phenomenon is an enhanced reliance on semantic processing to the detriment of the remembrance of punctual features of the items to be encoded (e.g., their position in a list, their size, the number of letters that compose them). On the other hand, it might be caused by the impairment in executive functioning that is typical of aging. We conducted an empirical study using DRM lists in two age groups, young and old, to disentangle which of the above mentioned two mechanisms might be responsible for the enhancement in verbal false memories found during aging. Our results support an unprecedented integration of these two views, which we discuss as an extension of the theoretical framework prominent in the literature.

Keywords: aging, false memory, DRM paradigm, semantics, executive functioning.

Introduction.

In everyday life, spontaneous false memories occur when someone believes to have experienced an event that did not really happen before. Under specific conditions, false memories can be elicited also in laboratory experiments, for instance, using the Deese-Roediger-McDermott paradigm (DRM paradigm, Deese, 1959; Roediger & McDermott, 1995). In the DRM task, participants are typically first asked to remember words presented sequentially; later they have to retrieve (i.e., recall or recognize) as many of the earlier encountered words as possible. The words shown to the participants are organized in semantically-related lists (e.g., drop, rain, canal, sea, mineral, sparkling, lake, faucet…), with all the terms in the lists being associated to a not present critical lure (e.g., water). Items are typically characterized according to the strength of their semantic association with the lure, from the most associated to the less associated (e.g., Beato & Arndt, 2017; Fam *et al.*, 2021).

At retrieval time, participants tend to produce some false memories: They may incorrectly indicate the presence of the critical lure (i.e.,, the word associated strongly with the theme of each list and therefore with the items actually presented). In the above example, this would be a word like 'water'. The DRM task has been applied for participants across the whole life span, yielding false memories in every age group (e.g., Sugrue & Hayne, 2006; Dennis *et al.*, 2007; Brainerd *et al.*, 2008;).

The occurrence of false memories has been explained as a consequence of semantic activation spreading from the list items to the lure by Fuzzy Trace Theory (FTT; Reyna *et al.*, 2016) and Spreading Activation Theory (SAT; Roediger *et al.*, 2001). The former theory postulates that item recollection occurs on the basis of two different and independent forms of processing. First, there is *verbatim* retrieval, which involves the elaboration of the surface characteristics of a trace to be memorized (e.g., the color of a stimulus or its position among other elements). Second, there is *gist* retrieval, referring to

the extrapolation of the meaning conveyed by what has been processed (e.g., Brainerd & Reyna, 2004; Reyna *et al.*, 2016; Howe & Derbish, 2010; Otgaar *et al.*, 2018). The latter theory posits that processing a concept results in activation spreading to related ideas in one's mind. This results in false remembrance, because the meaning of related words can be "incorrectly" activated.

The DRM has been argued to reliably measure semantic processing (Brainerd *et al.*, 2008). However, although a semantic contribution to false memory has been documented for younger adults and children (e.g., Brainerd *et al.*, 2009; Holliday *et al.*, 2011), the mechanism underlying false memories in older adults remains a topic of debate (e.g., Skinner & Fernandes, 2009). There is consistent empirical evidence that older adults show an increased number of verbal false memories relative to younger individuals (Devitt & Schacter, 2016; Jacoby & Rhodes, 2006; Norman & Schacter, 1997).

One potential reason for this finding is that older adults process semantic information in the stimulus list on the basis of general stimulus content (i.e., *gist* elaboration according to FTT, 2016), rather than on specific, punctual features as their position in a list, their size, the number of letters that compose them (i.e., *verbatim* processing according to FTT, 2016). According to this view, the increase of false memories with age may be traced back to a larger reliance on semantics, which is a domain that in itself is not affected by aging. However, another reason may be sought in the nature of the DRM task itself, which requires abilities that are known to decline in older adults, such as executive functioning (e.g., Murman, 2015).

In the present study, we aimed to clarify whether the increase of false memories as a function of age is dependent on a major reliance on semantic processing or rather on a diminished reliance on executive functioning. To examine this theoretical contrast in more detail, we will now briefly review available studies in support of each position. These two

explanations are apparently at odds. However, our purpose is to propose an unprecedented hypothesis that provided for a coexistence of both.

False memories' increase with age is related to changes in semantic processing

Regarding the semantic account, some authors have argued that the older they become, adults rely more and more on semantic elaboration when they have to memorize DRM lists. For instance, on the basis of three DRM-based experiments, Tun and colleagues (Tun *et al.*, 1998) proposed that the increased probability of false memory production in older adults can be understood as a consequence of increased reliance on a gist strategy. Similarly, Dennis, Kim, and Cabeza (2007) concluded that older adults are more prone to gist-based strategies and rely less on an item-specific encoding strategy. In their view, the larger engagement in gist-based processing is responsible for both true and false memories, as demonstrated by neural evidence showing an age-related activation of the superior temporal gyrus (STG), responsible for semantic processing (e.g., Visser & Ralph, 2011) and associated with both true and subsequent false retrieval, and not only with item-specific encoding. In a later study (2008), the authors suggested that the older participants' tendency to decline in true memory retrieval might reflect age-dependent changes to the hippocampus, implied in episodic memory and retrieval (e.g., Kramer *et al.*, 2005), while the enhanced reliance on semantic gist could be mediated by the lateral temporal cortex, known to be crucial for language comprehension (Goldstein *et al.*, 2017).

More generally, these findings appear to be consistent with studies assessing semantic processing during aging. On the basis of a review of studies, Birren (2006) concludes that older adults largely rely on their semantic knowledge when coding and retrieving laboratory-based or environmental information. Some studies even reported

better semantic performance for older adults when compared with younger adults. For example, Zhuang *et al.* (2016) investigated the neural underpinnings of semantic retrieval during a semantic similarity judgment task and found that the older adults had higher accuracy rates compared to the younger adults group.

Following up on these findings, Taylor and Burke (2002) presented phonologically related and semantically related distractors during a picture naming task and observed that only semantically related distractors interfered with the task. In their view, semantic memory capacity is preserved with age, both when considered on its own and when compared with episodic memory, which has been proven to decline with age.

In spite of this evidence, the view that the increased occurrence of false memories depends on an enhanced reliance in older adults on semantic memory has been challenged by other studies. For example, Au *et al*. (1995) and Barresi (2000) reported a memory decline at the semantic-lexical level starting from the age of 70. Similarly, Verhaegen and Poncelet (2012) administered the 'Pyrameeds and Palm Trees Test' (PPTT; Howard & Patterson, 1992), usually used to assess semantic access for words and pictures (e.g., Klein & Buchanan, 2009) and a 'Synonym and Judgment task' (Majerus *et al.*, 2001), usually used to assess the semantic knowledge (e.g., Antonucci *et al.*, 2008). The older adults were found to perform worse on these tests than young adult controls.

Thus, there is a body of evidence providing support for the second account on the increasing number of false memories across age, which holds that age difficulties are due to deteriorating executive functioning (e.g., Fisk & Sharp, 2004).

False memories' increase with age is related to changes in executive functioning

Askey and Playfoot (2018) explicitly contrasted a semantic and an inhibition hypothesis for the increase of false memories with age. First, they considered the Transmission Deficit Hypothesis, according to which there is a reduction of the strength of links in the semantic network due to aging. Second, the Inhibitory Deficit Hypothesis attributes the increase in false memory in aging to inhibition mechanisms that become less efficient. Based on the results of their experiments involving a manipulation of list length, the authors suggested that the Inhibition Deficit Hypothesis provides a better explanation for the formation of false memories in aging. In particular, this theory states that the semantic network works in older adults exactly as it does in younger adults, but during aging not necessary nodes get ignored not as efficiently. In addition, Askey and Playfoot criticized Tun et al.'s view (1998) of semantic enhancement in aging, arguing that an improvement in the semantic field should be followed not only by an increase in false memory but also in true memory, as a consequence of a general improvement in the task. However, from their study they did not find this pattern: they found an increase in verbal false memory and a reduction in true memory. Their result seems to indicate an impaired performance, instead of an enhanced one.

Aims of the study.

On these grounds, the main goal of this study is to systematically test the semantic and executive control accounts in the formation of the increased number of verbal false memories reported during aging. As a precondition for contrasting the two views, we expected to find a main effect of age in the production of false memories in line with the literature reviewed above: An increased number of falsely recognized critical lures should indeed be present in older adults. Thus, we first administered DRM lists to younger and older adults to assess their false memories' production.

Analyzing the literature, we expected both semantics and executive functioning to play a role in the increased production of verbal false memory during aging.

In fact, regarding the role of semantics, we expected that it might have more largely explained verbal false memories in the older adults in comparison with the younger adults. Our expectation was motivated by the literature affirming that semantic processing might be enhanced for the older adults (e.g., Birren, 2006).

To test our hypothesis, we computed the semantic similarity of the stimulus words by using the approach of Günther *et al*. (2019), defining Distributional Semantic Models as the category of models based on natural language data to represent word meanings as vectors. Furthermore, we followed the methodology used in Gatti *et al.* (in press), concerning how to use semantic similarity when considering verbal false memories. With respect to this dimension, we hypothesized that a higher semantic similarity of the words of the DRM lists would be correlated with a higher number of false memories, in line with recent findings (Gatti *et al.*, in press). We expected semantic similarity to account for our expectation.

With respect to the executive functioning, our hypothesis was that they may have different impacts depending on participant's age. In particular, we expected them to play a massive role in the older group, affecting their false memory performance. Our expectation was motivated by the fact that both inhibition and working memory are implicated in the recognition process and in the verbal false memory production (e.g., Lovden, 2003; Dodhia & MtCalfe, 1999; Peters et al., 2007) and by the fact that both these two aspects are known to decline due to aging (e.g., Birre, 2006).

We expected both semantics and executive functioning to be implicated, because our hypothesis was that inhibition may affect verbal false memories proportionally to the strength of the words' semantic similarity. Older adults may more hardly inhibit the more

semantically associated stimuli than the lower semantically associated ones. Indeed, our instance may account for an increased reliance on semantic processing, net of a proved decline in the executive functioning domain.

Method.

Participants

We included a total of 56 Italian monolingual participants, divided in two groups: one with 31 young participants (Mean Age = 28.69, Range = 25-32, Female participants = 64.52 %) and one with 25 older adults (Mean Age = 68.74, Range = 65-76, Female participants = 36 %).

As an inclusion criterion, participants in the younger group had to be between 20 and 35 years at the time of the testing and between 65 and 80 years for the older adults group. The group of older participants was administered the Mini Mental State Examination (MMSE, Folstein & McHugh, 1975), which assesses a person's neuro-cognitive and functional state. Only participants passing the 24-point *cutoff*, usually used as a benchmark for distinguishing healthy aging from pathological aging, were included in the older adults group. Because none of the older participants performed below threshold, all of them have been included in the sample.

Materials

All materials were presented, compiled, and collected via computer.

DRM Task.

We started from a pool of 32 DRM lists of 15 words each. The lists were adapted to Italian from the materials developed and validated by Brainerd and colleagues

(Cornell/Cortland Emotional Lists, CEL Lists, 2008). The validation of the adapted materials took place through a pilot study that involved 20 healthy older adults (Mean age: 74) and 20 young adults (Mean age: 26.5), who did not take part in the actual study. The frequencies of use of the words were considered referring to the ColFis Lexical Database (Bertinetto *et al*., 2005). The Backward Associative Strength index (BAS, i.e., how much each word included in the list is associated with the critical lure) was evaluated on a 7-point Likert scale (from 1 = not at all associated to 7 = very associated). Stimulus words were ordered from the most to the less associated on the basis of the assigned BAS derived from our pilot study on the Italian lexicon.

During the coding phase of the experiment, 12 of the 32 adapted lists were presented to participants, each containing the first 10 words mostly associated with the not present semantically associated critical lure. During the recognition phase, for each list we presented 3 words that were actually shown in the coding phase (9 in total), the critical lure relating to each list (12 in total), 1 semantically related distractor, that was the 12th word of the original lists of 15. The semantically related distractors were less correlated in terms of meaning than lures with the items in the list they belong to, but in any case fell within the same semantic sphere. i.e. the word "injection" from the list with "needle" as critical lure. To conclude, we inserted also 2 other distractors non related at the level of meaning, selected from the remaining 20 rejected lists of the 32 originally adapted (6 in total) (i.e., the word "spider").

Semantic similarity index.

We computed a semantic similarity index for the three categories of new words presented in the recognition phase: critical lures, semantically related distractors, and non-semantically related distractors. This index was computed by employing distributional semantic models (DSMs) that represent word meanings as high-

dimensional numerical vectors, extracted from large amounts of natural language data. DSMs, indeed, fulfil the criteria to be considered as psychological models of the nature of semantic representations and the structure of semantic memory (Günther, Rinaldi & Marelli, 2019).

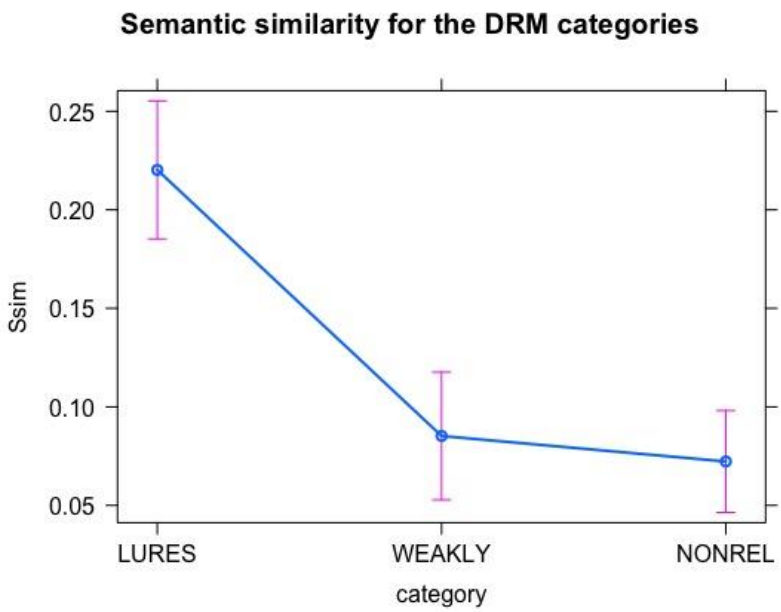
Vector representations for these words were extracted from the semantic space WEISS1 – Italian forms based on the Continuous Bag of Words (CBOW) model (Mikolov, Chen et al., 2013), trained on the Italian-text corpus ItWac. In order to determine the learning procedure to vectorize the words, parameters were set with 400 dimensions and 9-word window size. The window size indicates that predictions considered a target word in the middle of 4 words on the right and 4 words on the left; negative sampling $k = 10$ (i.e. "the probability of a target word by learning to distinguish it from draws from a noise distribution and with the parameter k that specifies the amount of these draws", Gatti *et al.*, 2021, *in press*) and subsampling with $t = 1e^{-5}$, corresponding to the threshold-based procedure aimed at limiting the impact of the high frequent uninformative words.

The words considered in the above mentioned semantic space were compared pairwise. For each pair, a semantic similarity index (SSim) was derived, expressed as the cosine of the vector angle obtained from the distance between the two words analyzed. The greater the cosine of the angle, the greater the proximity in terms of meaning between the two words. Each index obtained was subtracted from the value of 1, in order to rearrange the values on a proximity scale and weighted by its raw frequency (obtained by referring to the SUBTLEX-IT database, Crepaldi *et al*., 2016), as in Gatti *et al.* to the. (2021, *in press*).

To verify that the semantic similarity scores actually reflected differences in semantic relatedness with the studied words, we carried out a one-way ANOVA with the

semantic similarity index of the terms as a dependent variable, and the three categories of distractors as an independent categorical variable. The effect of word category was significant ($p < .01$, $F(2,85) = 25.58$ and 95% C.I.). In particular, the pairwise comparisons run on the categories showed a larger semantic similarity in the category of lures (Mean = .220, DS = .017, from estimated marginal means run on the one-way Anova), followed by the one of semantically related distractor types (Mean = .085, DS = .016, from estimated marginal means run on the one-way Anova) and, ultimately, by the category of non-semantically related distractors (Mean = .072, DS = .013, from estimated marginal means run on the one-way Anova) (see Figure 4.1).

Figure 4.1. Semantic Similarity for the DRM categories of lures, weakly related and unrelated.



The effect of semantic similarity on the number of false memories was also tested, with the categorization of distractors as covariate. This effect was significant ($p < .01$), proving that the semantic similarity index adds information over the simple tripartite categorization of distractors.

Cognitive Measures.

Inhibition task. We applied Hayling Sentence Completion Test (Burgess & Shallice, 1996). It consists of two sections, each composed of 15 sentences with the last word missing. In the first of the two sections (Section A), the participant is asked to complete a sentence with the correct word according to sentence context. An example of a sentence from Section A with correct completion is: “On a motorcycle you must always wear…” – helmet. In the second section (Section B), however, the participant is asked to complete the sentence with a word that is not related in terms of meaning to the context of the sentence, but which agrees by gender and number with the preceding article. An example of a sentence from Section B and its correct completion is: “Bees produce…” - traffic. Scoring was obtained considering accuracy for each sentence according to the guidelines included in the task. At the end of the two sections, a general index was obtained subtracting the scoring obtained in Section B to the scoring obtained in Section A.

Working Memory task. The Backward Digit-Span Task (Soylu, 2010) was run. The task was administered as a measure of working memory. The participant is verbally presented with series of numbers, divided into blocks of 3 sequences each. Within each block, the sequences have the same length. Blocks are divided taking into account the gradually increasing length of the sequences they contain.

Participants repeat each sequence, at the end of the oral presentation of the same sequence, in reverse order compared to how it was read (e.g. the experimenter reads: "5-7-9"; the participant replies: "9-7-5").

The task ends when the participant misses two out of three sequences within a block. The quantity of numbers of the last correctly executed block constitutes the score

obtained in the task. For example, participants will obtain a score of 4 if they miss two sequences in the 5-digit block.

Procedure

The test consisted of one individual session for each participant lasting about one hour, in telematic mode due to the Sars-Cov-2 (Covid-19) pandemic. The meeting took place via the Skype platform.

Initially, a first link to Google Forms was shared via e-mail. Via this link, consent to participate was obtained and some personal data were collected: age, education, and profession. The latter was measured as an index of socio-economic status through an adult adaptation of the Hollingshead Four Factor Index of Social Status questionnaire (1975). Participants were asked to share their screen, in order to read the form's consent lines together and to guide them to filling out the subsequent questions.

Before starting with the actual experiment, the Mini Mental State Examination (MMSE, Folstein & McHugh, 1975) was administered. The questions were read, and the answers marked directly by the experimenter on the sheet. Next, the original image of the associated drawing was shown on the screen, and participants were asked to copy their version on a white sheet of paper and show it clearly on screen. The image produced was acquired and stored via a screenshot.

The order of administration of the four tasks (three control tasks and the experimental task) was randomized across participants. Halfway through the meeting, a ten-minute break was scheduled.

DRM Task.

The experimental task was built through the OpenSesame 2.0 program (Mathôt, Schreij, & Theeuwes, 2012) and presented to the participants through the Just Another

Tool for Online Studies platform (JATOS, Lange, Kuhn & Filevich, 2015), which allowed sharing the task via link. The procedure adopted in the administration was the same as the one proposed by the authors of the original lists (Brainerd et al., 2008). During the coding phase, three lists of items appeared on the screen, at one word at a time, and with a 10 second gap between one list and the next. The stimuli were presented in "mono" font, size 18 in the center of the screen. Participants were instructed to memorize as many words as possible, because they would later be asked to recognize them among other words not previously seen. Following the first three item lists, the first recognition task was presented. Here the instruction was to press the space bar every time the participant recognized one of the previously memorized words, trying to be as fast and as accurate as possible. Between the coding phase and the recognition task, a pause of 30 seconds was provided, during which generic questions were asked relating to perceived fatigue or difficulty of the task. The aim of these questions was to distract the participants from the words just memorized. This procedure was repeated 4 times, each time with different lists.

Executive Functions.

The inhibition task and the working memory task were read aloud by the experimenter to participants. Answers were recorded and written directly by the experimenter while participants pronounced them.

Results.

For all analyses conducted below, we used the lme4 package and the lmerTest package in R studio 1.0.153 (RStudioTeam, 2020). This allowed us to build generalized linear mixed-effects models using the 'glmer' function.

We began by testing if aging was associated with an increased number of verbal false responses, in order to collect support for our confirmative hypothesis. For this purpose, we ran a first linear mixed model with the number of false memories (meaning the number of new words that were recognized as already seen) as a dependent variable, participant group (young vs. old) as a dichotomous independent variable, and participants and words as random intercepts.

To answer our main research question, we performed a model selection using the MuMIn R package, with the function 'dredge' (Bartoń, 2020). This procedure selects the best fitting model (i.e., the one with lowest Akaike information criterion, which returns an estimation of the quality of the model, AIC; Akaike, 1973) fitting all possible combinations of the fixed effects included.

We also checked whether one or more models had a Δ AIC < 2 compared with the best model identified by the dredge selection, as models in this AIC range can be considered as equivalent in explaining participants' performance (Hilbe, 2011). We only found one model with a Δ AIC=1.37, in which the interaction between group and span was dropped. However, in the full model reported in the Results section, the interaction between group and span was not significant. In light of this, we inserted false memory as dependent variable, group as main effect, in interaction with semantic similarity of the items, with inhibition (measured by the Hayling Test) and with working memory capacity (measured by the backward span test) (see Tables 4.1 and 4.2 for separate correlations between the covariates in the two groups).

Table 4.1. Correlation matrix between cognitive measures in the older adults group

	SSim	inhibition	Working memory
SSim	X	.00	.00

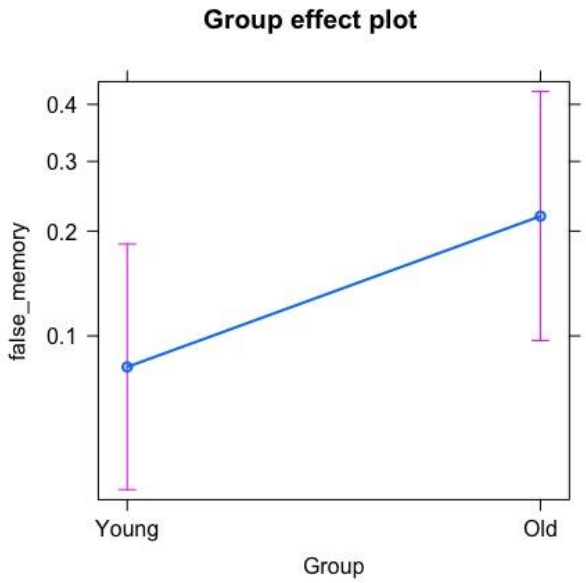
inhibition	0.0	X	-.08
Working memory	0.0	-.08	X

Table 4.2. Correlation matrix between cognitive measures in the younger adults group

	SSim	inhibition	Working memory
SSim	X	.00	.00
inhibition	0.0	X	-.02
Working memory	0.0	-.02	X

Concerning the first model (see Figure 4.2), built to test our confirmative hypothesis, the older adults group was found to produce a significantly higher number of "yes" responses for distractors relative to the young group (p = .031, Estimate = 1.167) .

Figure 4.2. False memory effect (the higher probability to respond "yes" to distractors) between the two groups

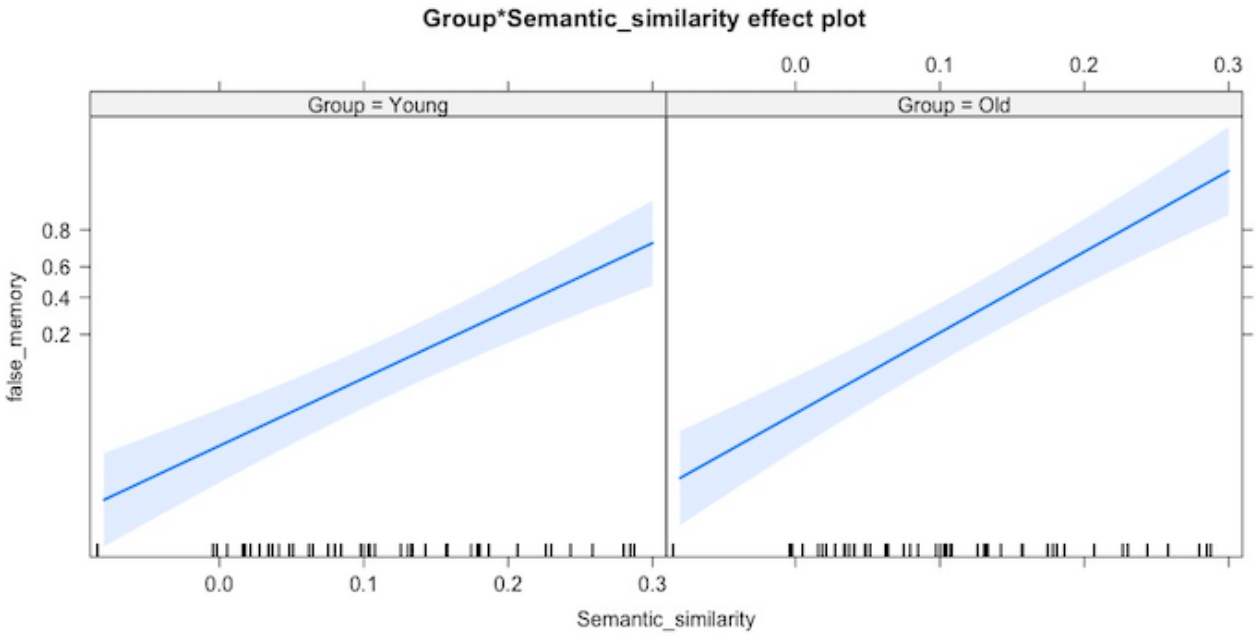


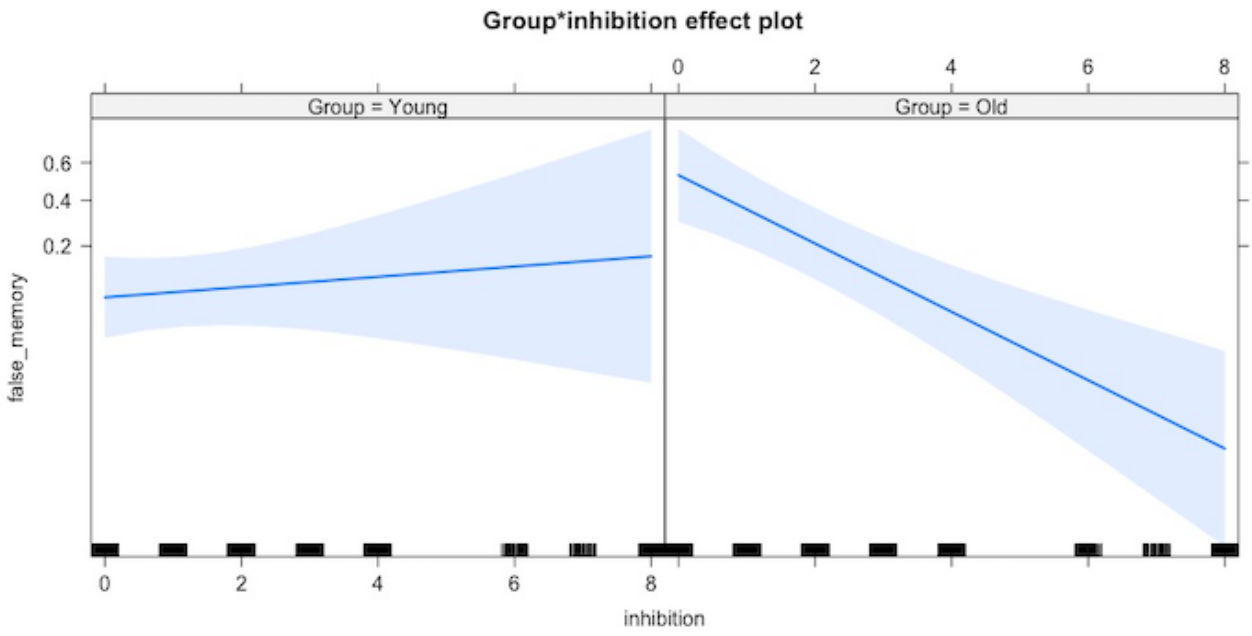
For the second model, aimed at answering our main research question, semantic similarity and group proved both to be significant (respectively with p <.001, t (46) = 9.130, Estimate = 17.891, $\chi 2$ (1) = 56.395 and p = .020, Estimate = 4.749, $\chi 2$ (1) =

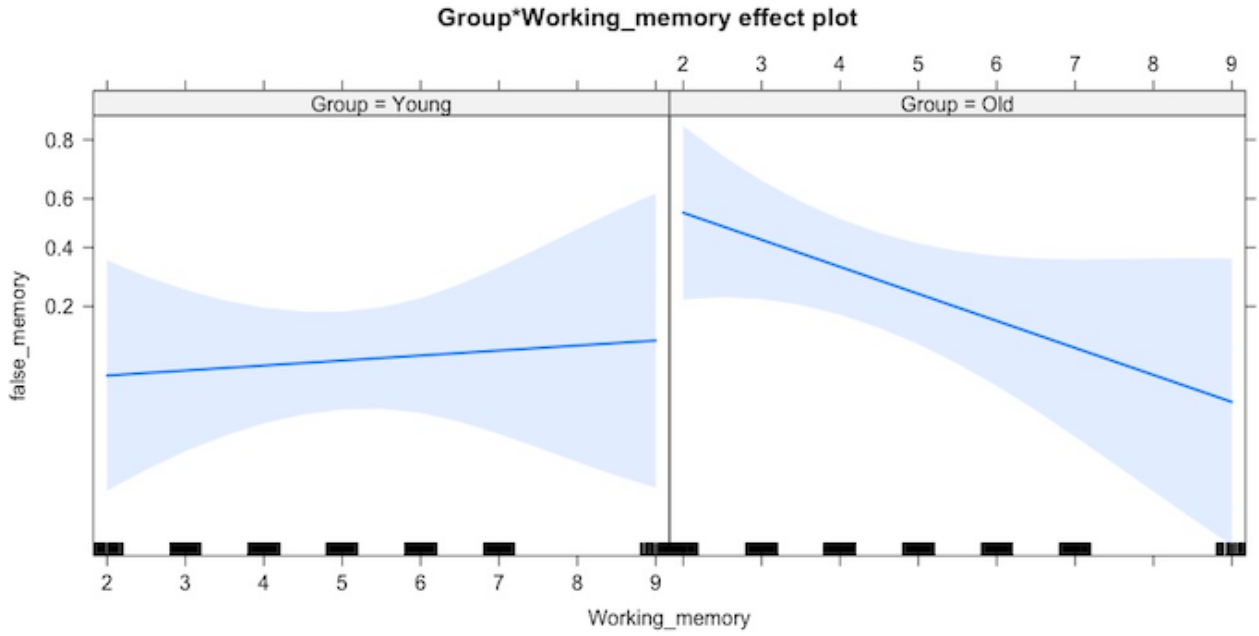
5.458). Inhibition significantly predicted the production of false memories in interaction with the older group ($p < .001$, Estimate = -.844, $\chi2$ (1) = 10.912). This demonstrates that a better performance in the inhibition task is accompanied by an increased production of false memories.

The interaction between semantic similarity and the older group showed a trend marginally significant ($p = .051$, Estimate = 3.450, $\chi2$ (1) = 3.808). In contrast, working memory was not significant in predicting false memory performance in the overall sample ($p = .780$, Estimate = 0.083, $\chi2$ (1) = .078), nor in interaction with the older group ($p = .173$, Estimate = - .533, $\chi2$ (1) = 1.856) (see Figure 4.3 for the interaction effects).

Figure 4.3.







Furthermore, we ran the same analyses after dividing (instead of multiplicating) the main effect (i.e., the variable "group") in terms of the other variables. The interaction with Semantic Similarity was significant in both groups (Older adults: $p < .001$, Estimate = 21.390, $\chi2$ (2) = 77.591; Younger adults: $p < .001$, Estimate = 17.891, $\chi2$ (2) = 77.591).

The interaction of group with the inhibition was significant only for the older adults ($p <$ .001, Estimate = - 0.734, $\chi2$ (2) = 20.001), but not for the younger adults (p = .571, Estimate = .110, $\chi2$ (2) = 20.001). The working memory measure did not significantly interact with either group (older adults: p = .077, Estimate = - .450, $\chi2$ (2) = 3.210; younger adults: p = .780, Estimate = .083, $\chi2$ (2) = 3.210).

Discussion.

In this study, we investigated whether the increased occurrence of false memories in older participants was due to a greater reliance on semantic processing (e.g., Tun et al., 1998; Dennis, Kim, and Cabeza, 2007; Dennis, Kim, and Cabeza, 2008) or to executive functioning impairment (e.g., Au et al., 1995; Barresi, 2000; Askey & Playfoot, 2018). Instead of contrasting these two aspects, we supposed that both could be correct and effective. To examine the former, we considered the semantic similarity between studied and new words from the DRM task, predicting performance in the same task based on this index. To consider the latter, we assessed the role played by inhibition and working memory, factors known to be implicated in false memories' production (e.g., Lovden, 2003; Dodhia & MtCalfe, 1999; Peters et al., 2007).

We first assessed our precondition, namely the increased number of verbal false memories in aging. In line with previous evidence (Devitt & Schacter, 2016; Jacoby & Rhodes, 2006; Norman & Schacter, 1997), older adults in our sample produced a higher number of false memories than younger participants.

With respect to semantics, we hypothesized that a) a higher semantic similarity of the words in the DRM lists would correlate with a higher number of false memories (e.g., Holliday et al., 2011; Brainerd et al., 2009); b) semantics might produce a larger effect on

verbal false memories in older adults than in younger adults (e.g., Birren, 2006, Tun *et al.*, 1998).

With respect to the former hypothesis, our results indicate that the semantic similarity of items significantly affected false memories' production in the general population, consisting of both younger and older participants. Despite the interaction between semantic similarity and older group tended to reach the significance, the analyses on the two age groups did not reveal a different pattern. The critical lures were the most falsely recognized items, followed by the semantically related distractors, and only then the semantically unrelated distractors.

These data do not allow us to consider semantics as the only responsible for the increased number of verbal false memories in older adults.

For the cognitive executive functioning measures, we had anticipated that they would have different impacts depending on participant group. In particular, we expected them to play a massive role in the older-adults group, since executive functioning is known to be affected by aging (e.g., Ullmann *et al*., 2020, Palumbo et al., 2018). Consequently, we expected executive factors to act more on the older adults' false memory performance than on the younger adults control group's false memory performance.

With respect to the older adults, we found that inhibition played a role, while working memory capacity was not significantly involved in the production of false memories. However unexpectedly inhibition showed a positive relationship with the production of false memories in older adults. In fact, although in this task the older adults performed overall worse than the younger adults (see Table 4.3), the older participants who obtained better results in terms of inhibition, produced more semantic false memories.

Table 4.3. T Test cognitive measures between groups

	Older Adults	Younger Adults	*p* value
Inhibition (Hayling Test)	2.185	1.103	< .01
Working Memory (Backward Span)	4.407	5.034	< .01

This result seems to go in the opposite direction of our hypothesis of worse inhibition for more semantically related not presented words. A possible explanation for this result could be the nature of the Hayling Test. Indeed, although it has been traditionally used as a measure of inhibition, there are also studies (e.g., Cervera-Crespo & Gonz lez-Alvarez, 2016) that used it as a measure of semantic control. With the term semantic control, we mean "the ability to selectively access and manipulate meaningful information on the basis of context demand" (Montefinese, Hallam, Thompson, & Jefferies, 2019; Jackson, 2021). Other studies reported a correlation between semantic control abilities and semantic representation (e.g., Chiou, Humphreys, Jung, & Ralph, 2018). It may be that those participants who were better in terms of semantic control, also produced more semantically related false memories. If on the one hand the lack of an inhibition measure detached from the verbal and semantic component can be considered as a limit of our study, on the other hand it lays the groundwork for considering semantic control as a possible innovative bridge between the two apparently contrasting theories that try to explain the increased number of verbal false memories in aging.

For a future study we aim to include other measures of inhibition (i.e., a measure of inhibition, a measure of suppression and a measure of restriction).

In conclusion, according to our hypothesis, we expected semantics to affect older adults' production of verbal false memories more than younger adults' one. In fact,

during aging the semantic knowledge not only does not decline, but rather it becomes the priority mechanism to rely on in the context of a generally declining memory performance and, in some cases, it even improves compared to young people. On the other hand, executive functioning is compromised in old age group (e.g., Birren, 2006). In light of this, we expected older people to have greater difficulty inhibiting the more semantically related items.

Our results seem only apparently to confirm initial hypotheses. Indeed, if on the one hand both semantics and executive functioning are involved in explaining the higher number of verbal false memories in older adults, on the other hand the latter plays an unexpected role. Indeed, older adults who produced more verbal false memories where the ones who performed better in the inhibition task. Given that it is a verbal task measuring the semantic control, a possible explanation of the association between inhibition and false memories could be that older adults need their executive functioning to be intact and strong in order to activate their semantic knowledge and the semantic network level that is implied in the DRM recognition task. According to this perspective, semantic activation in aging seems not to be the passive consequence of a difficult task that older adults need to manage with their compromised memory skills and their compromised executive functioning. Indeed, is seems to be a resource they can better activate if their executive functioning correctly works. Future studies should include different measure of inhibition to investigate better its role.

References.

Askey, C., & Playfoot, D. (2018). Examining theories of cognitive ageing using the false memory paradigm. *Quarterly Journal of Experimental Psychology*, *71*(4), 931–939. https://doi.org/10.1080/17470218.2017.1307433

Antonucci, S. M., Beeson, P. M., Labiner, D. M., & Rapcsak, S. Z. (2008). Lexical retrieval and semantic knowledge in patients with left inferior temporal lobe lesions. *Aphasiology*, *22*(3), 281–304. https://doi.org/10.1080/02687030701294491

Au, R., Joung, P., Nicholas, M., Obler, L.K., Kass, R. & Albert, M.L. (1995). Naming ability across the adult life span. *Aging and Cognition, 2*(4), 300-311. https://doi.org/10.1080/13825589508256605

Barresi, B. A., Nicholas, M., Tabor Connor, L., Obler, L. K., & Albert, M. L. (2000). Semantic Degradation and Lexical Access in Age-Related Naming Failures. *Aging, Neuropsychology, and Cognition*, *7*(3), 169–178. https://doi.org/10.1076/1382-5585(200009)7:3;1-q;ft169

Barton, K. (2020). MuMIn: Multi-model inference. https://CRAN.R-project.org/package=MuMIn Ben-Shachar

Beato, M. S., & Arndt, J. (2017). The role of backward associative strength in false recognition of DRM lists with multiple critical words. *Psicothema, 29*(3), 358-363.

Bertinetto, P. M. Burani, C. Laudanna, A. Marconi, L. Ratti, D. Rolando, C. (2005). Corpus e Lessico di Frequenza dell'Italiano Scritto (CoLFIS) http://www.istc.cnr.it/material/database/lexvar.htm.

Birren, J. E. *Handbook of the Psychology of Aging, Sixth Edition (Handbooks of Aging) (2005–12-23)*. (2021). Academic Press; 6 edition (2005–12-23).

Brainerd, C., & Reyna, V. (2004). Fuzzy-trace theory and memory development. *Developmental Review*, *24*(4), 396–439. https://doi.org/10.1016/j.dr.2004.08.005

Brainerd, C. J., Reyna, V. F., & Ceci, S. J. (2008). Developmental reversals in false memory: A review of data and theory. *Psychological Bulletin, 134*(3), 343–382. https://doi.org/10.1037/0033-2909.134.3.343

Brainerd, C. J., Yang, Y., Reyna, V. F., Howe, M. L., & Mills, B. A. (2008). Semantic processing in "associative" false memory. *Psychonomic Bulletin & Review*, *15*(6), 1035–1053. https://doi.org/10.3758/pbr.15.6.1035

Brainerd, C. J., Yang, Y., Toglia, M. P., Reyna, V. F., & Stahl, C. (2008). *Emotion and false memory: The Cornell/Cortland lists. In annual meeting of the Psychonomic Society*, Chicago, IL.

Burgess, P. W., & Shallice, T. (1996). Response suppression, initiation and strategy use following frontal lobe lesions. *Neuropsychologia*, *34*(4), 263–272. https://doi.org/10.1016/0028-3932(95)00104-2

Cervera-Crespo, T., & González-Alvarez, J. (2016). Age and Semantic Inhibition Measured by the Hayling Task: A Meta-Analysis. *Archives of Clinical Neuropsychology*. Published. https://doi.org/10.1093/arclin/acw088

Chiou, R., Humphreys, G. F., Jung, J., & Lambon Ralph, M. A. (2018). Controlled semantic cognition relies upon dynamic and flexible interactions between the executive 'semantic control' and hub-and-spoke 'semantic representation' systems. *Cortex*, *103*, 100–116. https://doi.org/10.1016/j.cortex.2018.02.018

Crepaldi, D., Amenta, S., Mandera, P., Keuleers, E., & Brysbaert, M. (2016). Frequency estimates from different registers explain different aspects of visual word recognition.

International Meeting of the Psychonomic Society, Granada, Spain, 5–8 May. http://crr.ugent.be/subtlex-it/

De Beni, R., & Palladino, P. (2000). Intrusion errors in working memory tasks. *Learning and Individual Differences*, *12*(2), 131–143. https://doi.org/10.1016/s1041-6080(01)00033-4

Deese, J. (1959) On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology, 58*, 17–22. DOI: 10.1037/h0046671

Dennis, N. A., Kim, H., & Cabeza, R. (2007). Effects of aging on true and false memory formation: An fMRI study. *Neuropsychologia*, *45*(14), 3157–3166. https://doi.org/10.1016/j.neuropsychologia.2007.07.003

Dennis, N. A., Kim, H., & Cabeza, R. (2008). Age-related Differences in Brain Activity during True and False Memory Retrieval. *Journal of Cognitive Neuroscience*, *20*(8), 1390–1402. https://doi.org/10.1162/jocn.2008.20096

Devitt, A. L., Tippett, L., Schacter, D. L., & Addis, D. R. (2016). Autobiographical memory conjunction errors in younger and older adults: Evidence for a role of inhibitory ability. *Psychology and Aging, 31*(8), 927–942. https://doi.org/10.1037/pag0000129

Dodhia, R. M., & Metcalfe, J. (1999). FALSE MEMORIES AND SOURCE MONITORING. *Cognitive Neuropsychology*, *16*(3–5), 489–508. https://doi.org/10.1080/026432999380898

Fam, J., Huff, M. J., Westbrook, R. F., & Holmes, N. M. (2021). The effect of early list manipulations on the DRM illusion. *Quarterly Journal of Experimental Psychology*, *74*(11), 1924–1934. https://doi.org/10.1177/17470218211012620

Fisk, J. E., & Sharp, C. A. (2004). Age-Related Impairment in Executive Functioning: Updating, Inhibition, Shifting, and Access. *Journal of Clinical and Experimental Neuropsychology*, *26*(7), 874–890. https://doi.org/10.1080/13803390490510680

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state." *Journal of Psychiatric Research*, *12*(3), 189–198. https://doi.org/10.1016/0022-3956(75)90026-6

Gatti, D., Rinaldi, L., Marelli, L., Mazzoni, G., Vecchi, T. (in press) Decomposing the semantic processes underpinning veridical and false memories. *Journal of Experimental Psychology*: General. IF 3.169

Goldstein, I. S., Erickson, D. J., Sleeper, L. A., Haynes, R. L., & Kinney, H. C. (2017). The Lateral Temporal Lobe in Early Human Life. *Journal of Neuropathology & Experimental Neurology*, *76*(6), 424–438. https://doi.org/10.1093/jnen/nlx026

Günther, F., Rinaldi, L., & Marelli, M. (2019). Vector-Space Models of Semantic Representation From a Cognitive Perspective: A Discussion of Common Misconceptions. *Perspectives on Psychological Science*, *14*(6), 1006–1033. https://doi.org/10.1177/1745691619861372

Henkel, L. A., Johnson, M. K., & De Leonardis, D. M. (1998). Aging and source monitoring: Cognitive processes and neuropsychological correlates. *Journal of Experimental Psychology: General, 127*(3), 251–268. https://doi.org/10.1037/0096-3445.127.3.251

Hilbe, J. M. (2011). Negative binomial regression. Cambridge University Press.

Holliday, R. E., Brainerd, C. J., & Reyna, V. F. (2011). Developmental reversals in false memory: Now you see them, now you don't! *Developmental Psychology*, *47*(2), 442–449. https://doi.org/10.1037/a0021058

Hollingshead, A. B. (1975). Four factor index of social status. Unpublished manuscript, Yale University, NewHaven, CT.

Howard, D., & Patterson, K. E. (1992). *The pyramids and palm trees test*. Bury St Edmunds, UK: The Thames Valley Test Company.

Howe, M. L., Wimmer, M. C., Gagnon, N., & Plumpton, S. (2009). An associative-activation theory of children's and adults' memory illusions. *Journal of Memory and Language*, *60*(2), 229–251. https://doi.org/10.1016/j.jml.2008.10.002

Howe, M. L., & Derbish, M. H. (2010). On the susceptibility of adaptive memory to false memory illusions. *Cognition*, *115*(2), 252–267. https://doi.org/10.1016/j.cognition.2009.12.016

Jackson, R. L. (2021). The neural correlates of semantic control revisited. *NeuroImage*, *224*, 117444. https://doi.org/10.1016/j.neuroimage.2020.117444

Jacoby, L. L., & Rhodes, M. G. (2006). False Remembering in the Aged. *Current Directions in Psychological Science*, *15*(2), 49–53. https://doi.org/10.1111/j.0963-7214.2006.00405.x

Johnson, M. K. (1988). Discriminating the origin of information. In T. F. Oltmanns & B. A. Maher (Eds.), *Delusional beliefs* (pp. 34–65). John Wiley & Sons.

Klein, L. A., & Buchanan, J. A. (2009). Psychometric properties of the Pyramids and Palm Trees Test. *Journal of Clinical and Experimental Neuropsychology*, *31*(7), 803–808. https://doi.org/10.1080/13803390802508926

Kramer, J. H., Rosen, H. J., Du, A.-T., Schuff, N., Hollnagel, C., Weiner, M. W., Miller, B. L., & Delis, D. C. (2005). Dissociations in Hippocampal and Frontal Contributions to Episodic Memory Performance. *Neuropsychology, 19*(6), 799–805. https://doi.org/10.1037/0894-4105.19.6.799

Lange K, Kühn S, Filevich E (2015) Correction: "Just Another Tool for Online Studies" (JATOS): An Easy Solution for Setup and Management of Web Servers Supporting Online Studies. PLOS ONE 10(7): e0134073. https://doi.org/10.1371/journal.pone.0134073

Lecce, S., Zocchi, S., Pagnin, A., Palladino, P., & Taumoepeau, M. (2010). Reading Minds: The Relation Between Children's Mental State Knowledge and Their Metaknowledge About Reading. *Child Development*, *81*(6), 1876–1893. https://doi.org/10.1111/j.1467-8624.2010.01516.x

Loebner, S. (2013). Understanding Semantics. *Taylor & Francis Group*. https://doi.org/10.4324/9780203528334

Lovden, M. (2003). The episodic memory and inhibition accounts of age-related increases in false memories: A consistency check*1. *Journal of Memory and Language*, *49*(2), 268–283. https://doi.org/10.1016/s0749-596x(03)00069-x

Majerus, S., Lekeu, F., Van der Linden, M., & Salmon, E. (2001). Deep dysphasia: Further evidence on the relationship between phonological short-term memory and language processing impairments. *Cognitive Neuropsychology, 18(5)*, 385–410.

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. doi:10.3758/s13428-011-0168-7

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. Retrieved from https://arxiv.org/abs/1301.3781.

Mitchell, K.J., & Johnson, M.K. (2000). Source monitoring: Attributing mental experiences. In E. Tulving & F.I.M. Craik (Eds.), *The Oxford handbook of memory* (pp. 179-195). New York: Oxford University Press.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The Unity and Diversity of Executive Functions and Their Contributions to Complex "Frontal Lobe" Tasks: A Latent Variable Analysis. *Cognitive Psychology*, *41*(1), 49–100. https://doi.org/10.1006/cogp.1999.0734

Montefinese, M., Hallam, G., Thompson, H. E., & Jefferies, E. (2019). The interplay between control processes and feature relevance: Evidence from dual-task methodology. *Quarterly Journal of Experimental Psychology*, *73*(3), 384–395. https://doi.org/10.1177/1747021819877163

Murman, D. (2015). The Impact of Age on Cognition. *Seminars in Hearing*, *36*(03), 111–121. https://doi.org/10.1055/s-0035-1555115

Norman, K. A., & Schacter, D. L. (1997). False recognition in younger and older adults: Exploring the characteristics of illusory memories. *Memory & Cognition, 25*(6), 838–848. https://doi.org/10.3758/BF03211328

Otgaar, H., Howe, M. L., Muris, P., & Merckelbach, H. (2018). Associative Activation as a Mechanism Underlying False Memory Formation. *Clinical Psychological Science*, *7*(2), 191–195. https://doi.org/10.1177/2167702618807189

Otgaar, H., Howe, M. L., Muris, P., & Merckelbach, H. (2019). Dealing With False Memories in Children and Adults: Recommendations for the Legal Arena. *Policy Insights from the Behavioral and Brain Sciences*, *6*(1), 87–93. https://doi.org/10.1177/2372732218818584

Palmer, F. R. (2021). *Semantics by Frank Robert Palmer (1981–10-30)*. Cambridge University Press.

Palumbo, R., Mammarella, N., di Domenico, A., & Fairfield, B. (2018). When and where in aging: the role of music on source monitoring. *Aging Clinical and Experimental Research*, *30*(6), 669–676. https://doi.org/10.1007/s40520-018-0955-4

Peters, M. J., Jelicic, M., Verbeek, H., & Merckelbach, H. (2007). Poor working memory predicts false memories. *European Journal of Cognitive Psychology, 19(2*), 213-232.

RStudio Team (2020). RStudio: Integrated Development for R. *RStudio, PBC, Boston, MA URL http://www.rstudio.com/*

Reyna, V. F., Corbin, J. C., Weldon, R. B., & Brainerd, C. J. (2016). How fuzzy-trace theory predicts true and false memories for words, sentences, and narratives. *Journal of Applied Research in Memory and Cognition*, *5*(1), 1–9. https://doi.org/10.1016/j.jarmac.2015.12.003

Roediger, H. L., III McDermott, K. B. (1995). Creating false memories: Remembering words not presented in list. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 803–814. DOI: 10.1037/0278-7393.21.4.803

Skinner, E. I., & Fernandes, M. A. (2009). Illusory recollection in older adults and younger adults under divided attention. *Psychology and Aging, 24*(1), 211–216. https://doi.org/10.1037/a0014177

Soylu, F. (2010). Forward/backward digit-span task. *Arch. Neurobehav. Exp. Stimuli* 222. http://www.neurobs.com/ex_files/expt_view?id=218

Sugrue, K., & Hayne, H. (2006). False memories produced by children and adults in the DRM paradigm. *Applied Cognitive Psychology*, *20*(5), 625–631. https://doi.org/10.1002/acp.1214

Taylor, J. K., & Burke, D. M. (2002). Asymmetric aging effects on semantic and phonological processes: Naming in the picture-word interference task. *Psychology and Aging, 17*(4), 662–676. https://doi.org/10.1037/0882-7974.17.4.662

Tun, P. A., Wingfield, A., Rosen, M. J., & Blanchard, L. (1998). Response latencies for false memories: Gist-based processes in normal aging. *Psychology and Aging*, *13*(2), 230–241. https://doi.org/10.1037/0882-7974.13.2.230

Ullmann, G., Li, Y., Ray, M. A., & Lee, S. T. (2020). Study protocol of a randomized intervention study to explore effects of a pure physical training and a mind–body exercise on cognitive executive function in independent living adults age 65–85. *Aging Clinical and Experimental Research, 33*(5), 1259–1266. https://doi.org/10.1007/s40520-020-01633-w

Verhaegen, C., & Poncelet, M. (2012). Changes in Naming and Semantic Abilities With Aging From 50 to 90 years. *Journal of the International Neuropsychological Society*, *19*(2), 119–126. https://doi.org/10.1017/s1355617712001178

Visser, M., & Lambon Ralph, M. A. (2011). Differential Contributions of Bilateral Ventral Anterior Temporal Lobe and Left Anterior Superior Temporal Gyrus to Semantic Processes. *Journal of Cognitive Neuroscience*, *23*(10), 3121–3131. https://doi.org/10.1162/jocn_a_00007

Yee, E., Chrysikou, E. G., & Thompson-Schill, S. L. (2014). Semantic memory. In K. N. Ochsner & S. M. Kosslyn (Eds.), *The Oxford handbook of cognitive neuroscience, Vol. 1. Core topics* (pp. 353–374). Oxford University Press. https://psycnet.apa.org/record/2014-16124-017

Zhuang, J., Johnson, M. A., Madden, D. J., Burke, D. M., & Diaz, M. T. (2016). Age-related differences in resolving semantic and phonological competition during receptive language tasks. *Neuropsychologia*, *93*, 189–199. https://doi.org/10.1016/j.neuropsychologia.2016.10.016

In the previous chapter we aimed to address why older adults produce more verbal false memories than younger adults. In particular, we aimed to disambiguate a debate that we found in the literature between studies who considered this increase in terms of a greater reliance on semantic processing and other studies that considered it in terms of mnestic errors that are the consequence of a general decline in terms of memory and of executive functioning. Our results allowed us to support an integration between these two views, but in an unexpected direction. Indeed, semantics seems to be related to some executive functioning domain. In particular, a good inhibition functioning seems to determine a good level of the kind of semantic processing that is implied in the DRM false memory paradigm. For the purposes of the previous study we decided to momentarily set aside the role of emotionality and consider our research question as a general precondition. Starting from our results, in the next chapter we are going to take one step further, considering in an exploratory way the role that emotionality might play in verbal false memories' production during aging.

CHAPTER 6

The role of emotionality on verbal false memories in older adults.

Introduction.

Verbal false memories have been traditionally measured by means of the DRM paradigm (Deese, 1959; Roediger & McDermott, 1995). Participants are presented with lists of words and asked to memorize as many of them as possible. Every list is composed by words all semantically related to a them that is not part of the list (called "critical lure"). Participants are likely to remember also critical lures as if they actually encoded them. This phenomenon occurs on the basis of a tendency to process information referring more on its gist than on its punctual and specific features, as it has been stated by Fuzzy Trace Theory (FTT; Reyna *et al.*, 2016). Furthermore, there is a spreading of activation while processing the words that reinforce also the concept that is conveyed by critical lures, as stated by Spreading Activation Theory (SAT; Roediger *et al.*, 2001).

Numerous studies on monolingual adults addressed that the higher emotionality of the words would be associated to a larger number of verbal false memories. This would be because the emotional connotation of the words (namely, the emotional component connected to the meaning of the words) lead to process those words referring even more to semantics, in light of the priority that they acquire in our minds (e.g., Zhang, 2016).

Numerous studies demonstrated that high arousal words (i.e., more emotional words) would be associated to a significantly higher number of verbal false memories (e.g., , Sharkawy, Groth, Vetter, Beraldi, & Fast, 2008; Dehon, Larøi, & Van der Linden, 2010). Instead, with respect to valence, although there is less unanimity between studies (e.g., Dehon, Larøi & Van der Linden, 2010), it has been addressed that negative valence words would be associated with a higher number of verbal false memories (e.g., Brainerd, Stein, Silveira, Rohenkohol, & Reyna, 2008).

As far as we reviewed, only Kensinger & Corkin (2004) considered the effect of emotionality on verbal false memories in aging. They administered DRM lists with lures that were divided in neutral ones and emotional ones to both older adults and a younger control group. Their hypothesis was that emotionality would provide distinctiveness to encoding and, though, protect against false memories for critical lures. Their results confirmed this position. In fact, emotional lures were less falsely recalled or recognized than the neutral ones in both the age groups, the older adults one and the younger adults one. It has to be noticed that valence has not been properly included in their experimental design. Their results go in the opposite direction of the above mentioned literature on the other age groups (e.g., Sharkawy, Groth, Vetter, Beraldi, & Fast, 2008; Dehon, Larøi, & Van der Linden, 2010), both concerning their experimental older adults groups and also concerning their younger adults control group.

With respect to the effect of valence on verbal false memory during aging, there is a limited number of studies, which only considered valence per se, not including its interaction with arousal.

In fact, both, Piguet, Connally, Krendl, Huot, and Corkin (2008) and Fernandes, Ross, Wiegand, and Schryer (2008) tested the effect of valence during aging on false memory and found that positive words enhance verbal false memory for critical lures, in opposition with results found on younger adults (e.g., e.g., Brainerd, Stein, Silveira, Rohenkohol, & Reyna, 2008). This effect could be due to the bias of positivity that has been found for older adults. With this term we refer to the age-related trend to favor and focus more on positive items over negative ones (e.g., Reed & Carstensen, 2012). Although, in Piguet and colleagues' results (2008) positive stimuli were also associated with a better correct rejection of unstudied items, less semantically correlated to the encoded ones. As a whole, these results seem to be consistent with Zhang *et al.* (2016)

review, according to which emotionality would act a booster for semantic processing, which is responsible for false memory of critical lures and protective when this semantics would not be so prominent.

In light of Kensinger and Corkin's results' discrepancy with respect to other literature concerning arousal, we decided to further explore the role of emotion on verbal false memories during aging, also including valence in interaction with arousal in our experimental design. In fact, these two variables have not properly been tested together, as it has already been done on other age groups (e.g., Brainerd *et al.*, 2010). Indeed, we think that this interaction could provide insights on the role that the emotionality of the words has on semantic processing in aging.

Aims of the study.

Aim of the present study is to assess how the emotional component of the lexicon (represent both by their arousal and by their valence) would affect older adults' verbal false memories.

We know from the literature that older adults tend to produce a larger number of verbal false memories than the other age groups (e.g., Devitt & Schacter, 2016; Jacoby & Rhodes, 2006; Norman & Schacter, 1997). Our previous study demonstrated that semantic processing is at least partially involved in this increase (for further details please refer to the previous study).

We also know that emotionality proved to act as a booster for semantic processing in the other age populations, being responsible for a larger number of verbal false memories associated to higher emotionally charged critical lures. We would like to address if emotionality affects older adults' verbal false memories' production similarly to the way in which it does in other age groups. Kensinger and Corkin (2004) found that

emotion (in their study considered only as arousal and lacking the valence component) would diminish verbal false memories, both in their older adults experimental groups and in their younger adults control group.

Our hypothesis is that older adults' verbal false memories production might be affected by arousal similarly to younger adults, consistently with findings in more recent literature (e.g., Sharkawy, Groth, Vetter, Beraldi, & Fast, 2008; Dehon, Larøi, & Van der Linden, 2010). More in detail, we expect higher arousal lures to be more falsely remembered than lower arousal lures.

Arousal seems to play a similar role on memory in both younger and older adults (e.g., Sutherland & Mather, 2015) and the response it activate also seems not to decline during aging (e.g., Kunzmann, Kupperbusch, & Levenson, 2005; Otani, Libkuman, Widner, & Graves, 2010), leading us to believe that it might affect verbal false memory similarly to what it does on younger adults.

Concerning valence, we expect that older adults might be differently affected by valence in light of the bias of positivity that has been found for older adults. (e.g., Reed & Carstensen, 2012). More in detail, while we expect younger adults to produce more verbal false memories for negative critical lures, with respect to older adults we expect them to show the opposite pattern. This expectation is based on Fernandes and colleagues' results (2008) and on Piguet and colleagues results (2008).

Method.

With respect to the method section, please refer to the previous study for more details.

Participants.

Participants were the same of the previously described study. We included a total of 56 monolingual Italian participants, divided in two groups: one with 31 young participants (Mean Age = 28.69, Range = 25-32, Female participants = 64.52 %) and one with 25 older adults (Mean Age = 68.74, Range = 65-76, Female participants = 36 %).

Materials.

Experimental task.

Cornell/Cortland Emotional DRM Lists (Brainerd et al., 2008).

In order to measure verbal false memory, our participants were exposed to 12 emotional DRM lists. The materials have been adapted and validated from the original Cornell/Cortland Emotional lists (CEL Lists, 2008), following the same testing procedure used by the Authors with adults.

Every list presented in the encoding phase was made up of 10 of the 15 original words. The 12 lists were 3 from each of the four Arousal * Valence combination: High Arousal * Negative Valence, Low Arousal * Negative Valence, High Arousal * Positive Valence, Low Arousal * Positive Valence.

The recognition task was related to every Arousal * Valence combination and was made up of 21 words: 9 target words (3 from each presented list), the 3 not presented critical lures corresponding to each presented list, 1 related distractor for each presented list (taken by the 12th position of the 15-words lists) and 6 unrelated distractors, two for each of three unpresented lists, belonging to the same Arousal * Valence combination.

Control measures.

As control measures, we used The same measures that we used as additional measures in the previous paper, adding here two vocabulary tasks, consistently with the design that have been adopted with other populations.

Hayling Sentence Completion Test (Burgess & Shallice, 1996): we inserted this measure to consider inhibition as a control variable.

Forward/Backward Digit-Span Task (Soylu, 2010): both the versions of the task were administered as a measure of short-term memory.

PMA Synonym Task (PMA; 1981) and Categorical Fluency Task: these two tasks have been inserted as an overall index to measure vocabulary.

In the former, participants had to choose between four possibilities which was the correct synonym of a given word. The items were eighteen. Participants had 3 minutes to complete the task. They could skip the items that they didn't know.

In the latter, participants had to name as many words as possible belonging to a give category (i.e., "animals", "car brands" and "fruits"). They had 1 minute for every category. Scores were obtained considering the total correct words produced for each category in the given time, as in Krumm and colleagues (2021).

Center for Epidemiological Studies Depression Scale (CES-D; Radloff, 1977): In addition to the previously exposed measures, we included also this questionnaire, composed by 20 items asking participants to report how often they experienced symptoms associated with depression over the past week (i.e., sense of loneliness, loss of appetite, difficulty in sleeping). Answer could be between 0 and 3, where 0 was equal to None of the Time, while 3 was equal to Most or Almost all the Time.

Procedure.

The procedure adopted was the same of the previous study. Participants took part in a telematic meeting lasting about one hour trough the Skype platform. They had to sign the informed consent to take part in the study via Google form shared by the experimenter by chat. After signing, they had to answer the CES-D questionnaire through the same Google form. While completing the CES-D questionnaire, they had to share their screen, allowing the experimenter to guide them during the compilation or answering to possible questions.

Experimental task.

The DRM task was built through OpenSesame 2.0 program (Mathôt, Schreij, & Theeuwes, 2012) and proposed through the Just Another Tool for Online Studies platform (JATOS, Lange, Kuhn & Filevich, 2015). The experimenter shared the link obtained through JATOS by chat. Participants had to share their screen while completing the task. After reading the instructions, there was a shirt practice made by neutral DRM lists to be decoded before the real experiment started.

The procedure adopted in the administration was the same proposed by the authors of the original lists (Brainerd et al., 2008). It consisted of an encoding phase consisting on three lists belonging to the same arousal * valence condition. After a short delay phase there was the first recognition task, referring to the three coded lists. Words were present on the screen one at a time. They were written in “mono” font, in the center of the screen. During the recognition phase participants had to press space every time they thought they recognized one of the actually coded words. This procedure was repeated for every of the four arousal*valence blocks.

Control measures.

The administration of the control measures was randomized between participants.

Vocabulary and Categorical Fluency

The vocabulary task was administered as a Google Form link. Participants had to share their screen while completing the task. The instructions were read aloud together with the participant, accompanied by an example before starting the task.

With respect to the Categorical Fluency Task, the instructions were given aloud by the experimenter, providing an example of words belonging to a category that was not included in the real task. Words were written by the experimenter on a word document.

Forward/Backward Digit-Span Task

Both for the Forward and for the Backward version of the Digit-Span Task, the experimenter read the sequences of numbers aloud, writing down the participant's answers. The sequences were read at a rate of approximately 1 second for every number. A trial was provided as example before starting the testing for every of the two tasks.

CES-D Questionnaire

With respect to the CES-D Questionnaire, instructions were given by the experimenter, then participants were free to answer without sharing their screen in order to respect their privacy. They were encouraged to ask for questions at any time.

Results.

Data planning

All analyses were done using the package "lme4" in R Studio 1.0.153 (RStudioTeam, 2020) for mixed effect models.

As a dependent variable we referred to d prime index, as in the previous emotional DRM studies.

d prime is the index used to separate memory performance from noise (i.e., response bias). When considering verbal false memory, 'hits' usually correspond to the correctly recognized words, while 'false alarms' usually correspond to critical lures. The d' indexes obtained (z(H) – z(FA)) were the ones considered in the analyses, so as to work on a clean measure (Milfont & Fischer, 2010). (See Figure 5.1)

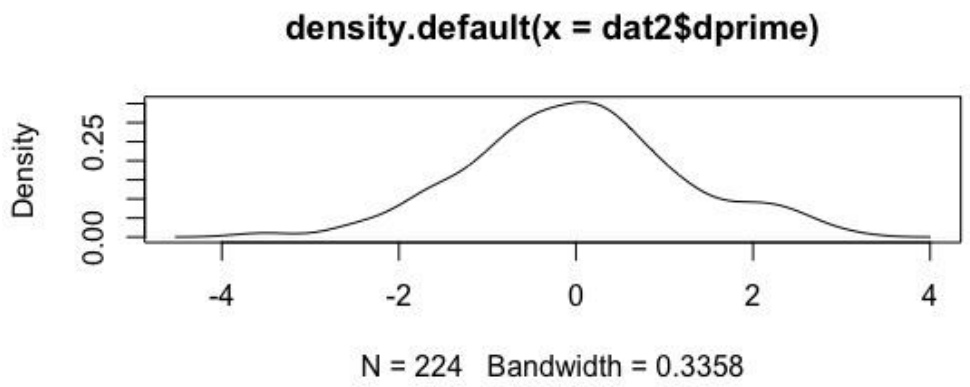


Figure 5.1. d prime values' distribution

As main factors we inserted Group in interaction with valence and arousal.

Group was considered as a dichotomous variable coded as 0 (i.e., Young adults) – 1 (i.e., Older adults).

Valence and arousal were considered as dichotomous too, the former as negative or positive, the latter as high or low. As covariates we inserted Span, Vocabulary, Inhibition and CES-D[2]. Span was intended as an overall index resulting from the average between Direct and Backward Span results. Similarly, Vocabulary was inserted as an overall index resulting from the average between PMA results and Categorical Fluency results. We did this referring to Lecce, Zocchi, Pagnin, Palladino & Taumoepeau (2010). Mixed models allow you to insert random effects. Here we considered as random effects arousal * valence with respect to participants. In this way we considered words (here represented by their emotionality) and participants variability.

Analysis

The results related to verbal false memory production remain the same described in the previous study. Group is still significant in affecting verbal false memory (p = .002, $\chi 2$ (1) = 6.650) also considering d prime, with an estimate = -1.066, meaning that being old is associated to lower *d* prime values, that correspond to higher false memories. Inhibition remains significant (*p* = .04, $\chi 2$ (1) = 7.654), as already assessed and discussed in the previous study, with an Estimate = -.114, meaning that higher scores obtained in the Hayling task correspond to lower *d* prime values, namely a larger number of false memories.

With respect to the newly considered dimensions, Vocabulary turned out to be significant (*p* = .009, $\chi 2$ (1) = 1.029) in affecting verbal false memories, with an estimate

[2] We tried to verify the role of CES D questionnaire in interaction with the emotional variables, but it turned out not to be significant, neither in interaction with arousal (*p* = .746, $\chi 2$ (2) = 2.676), neither in interaction with valence (*p* = .492, $\chi 2$ (1) = 1.088), neither in a three-way interaction with both arousal and valence (*p* = .362, $\chi 2$ (1) = .822).

= .181, meaning that better vocabulary results would be associated to higher accuracy rates both in the group of younger and older adults.

With respect to the emotional pattern, no differences has been found between the two groups, as demonstrated by the absence of significance in the group * arousal interaction (p = .154, Estimate = .467, $\chi2$ (2) = 2.104), in the group * valence interaction (p = .711, Estimate = .118, $\chi2$ (1) = .739) and in the group * valence * arousal interaction (p = .241, Estimate = - .443, $\chi2$ (1) = 1.145).

Low arousal in interaction with positive valence turned out to have a trend towards significance (p = .070, $\chi2$ (1) = 1.311), with an Estimate = .981, meaning that in both the groups low arousal and positive valence seem to go in the direction of a reduction of the number of verbal false memory, consistently with literature. Span and CES-D were not significant (respectively, p = .135, Estimate = .155;, $\chi2$ (1) = 2.027, p = .387, Estimate = .013, $\chi2$ (1) = 1.740). (See Figure 5.2 and Figure 5.3)

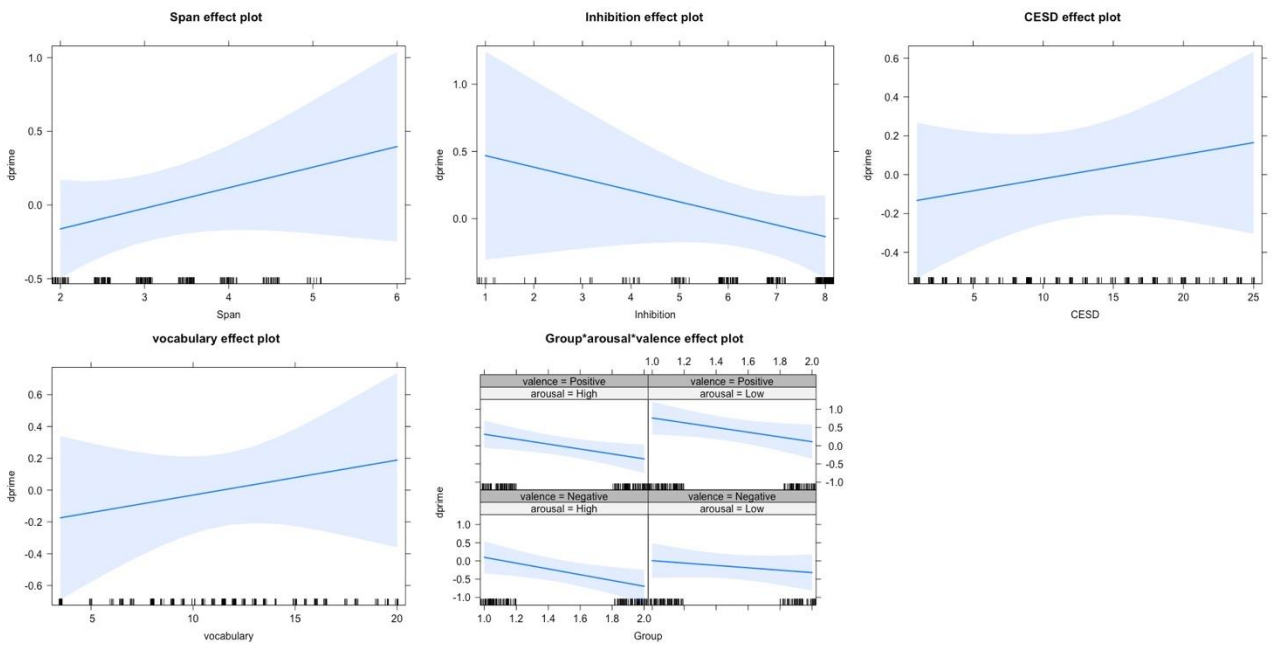


Figure 5.2. Mixed Effect Model

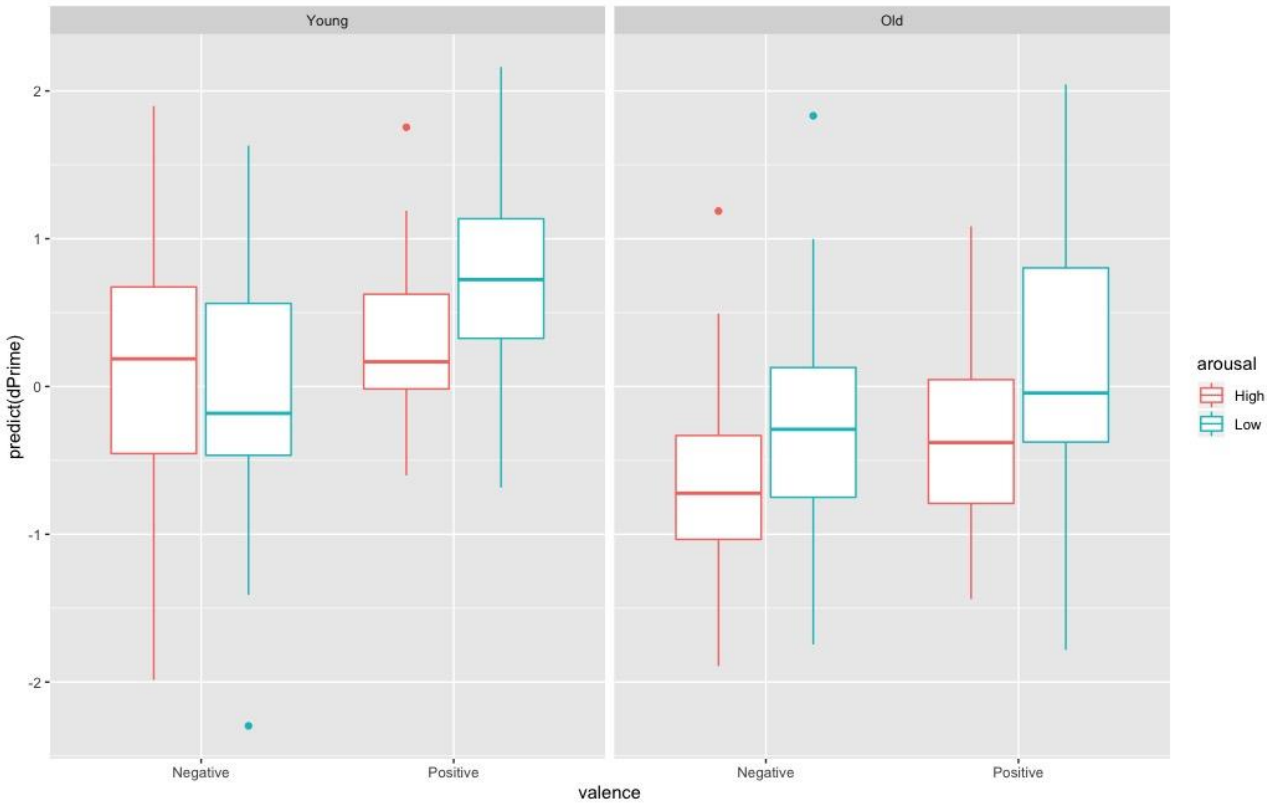


Figure 5.3. Emotion effect

Discussion.

The present study aimed to assess the influence of the emotionality of the lexicon on older adults' verbal false memories. In particular, we aimed to verify if arousal and valence of semantically related DRM critical lures could predict differences in verbal false memories, as it was found on the other age groups (e.g., Sharkawy, Groth, Vetter, Beraldi, & Fast, 2008; Dehon, Larøi, & Van der Linden, 2010).

Although our results were not strictly significant and need further exploration, we can discuss them and the direction of their effects both with respect to false memory and with respect to accuracy.

As far as we reviewed, there is only one study concerning arousal effect on verbal false memories during aging (Kensinger & Corkin, 2004). The authors found that the

higher emotionality of the DRM critical lures seems to be significant in reducing the number of verbal false memories both in their older adults experimental group and in their younger adults control group. They explained their results in light of the distinctiveness provided by the emotional value. This result is in opposition with the ones found in other similar studies concerning the role of emotionality on verbal false memories as measured on younger adults with an age similar to that of their control group (e.g., Sharkawy, Groth, Vetter, Beraldi, & Fast, 2008).

In light of this discrepancy, we decided to further explore emotional verbal false memories in aging, considering also valence in our experimental design, a dimension that has not been included in Kensinger and Corkin's study (2004).

We had two hypothesis, one concerning the effect of arousal and the other concerning the effect of valence.

With respect to the former, we expected higher arousal lures to be more falsely remembered than lower arousal lures in both the age groups.

This hypothesis was motivated by the above mentioned literature that assessed the role of arousal in younger adults and that found that arousal acts as a booster for semantic processing, leading to a higher degree of distortion and, consequently to a higher number of verbal false memories (e.g, Zhang, 2016). Some literature addressed that arousal plays a comparable role on memory performance in both younger and older adults (e.g., Sutherland & Mather, 2015; Otani *et al.*, 2005) and a similar response to it (e.g., Kunzmann *et al.*, 2005). These results lead us to believe that it might affect verbal false memory similarly to what it does on younger adults.

Coherently with our hypothesis and with the literature, we found no differences between the two groups in the way in which arousal affect verbal false memories. Even if this result is not significant, it is possible to notice from the plot (Figure 3) that high

arousal lures are generally more often falsely remembered than low arousal lures. The only exception comes from the condition of negative valence in the younger group, that seems to cause a higher number of false memories in the condition of low arousal than in the condition of high arousal.

With respect to valence, we expected it to be differently affected by aging in light of the bias of positivity that has been found for older adults (e.g., Reed & Carstensen, 2012), meaning with this term the age-related trend to favor and focus more on positive items over negative ones (e.g., Reed & Carstensen, 2012). More in detail, while we expected younger adults to produce more verbal false memories for negative critical lures, we expected older adults we to falsely remember a larger number of positive critical lures. These hypothesis was based on Fernandes and colleagues (2008) and on Piguet and colleagues (2008) results, who both studied the influence that valence has on false memory in aging and found that positive items elicit a larger number of false memories in the older adults for critical lures, while younger adults falsely remembered negative items more often than positive ones. Concerning older adults, positive items also increased correct rejection for less semantically prominent distractors.

We found no significant differences in the way in which valence affected verbal false memories between our two groups. This result seems to be in contrast with the previous literature and with our hypothesis, according to which we expected to find older adults producing the higher number of false memories for positive valence lures and with younger adults producing the higher number of false memories for negative items.

We found a trend towards significance for the two way interaction between low arousal and positive valence. In fact, this condition seems to increase accuracy in both the groups. This result is consistent with the previous literature concerning the effect of emotionality on younger adults (e.g., Sharkawy, Groth, Vetter, Beraldi, & Fast, 2008;

Dehon, Larøi, & Van der Linden, 2010). Concerning our control measures (namely, vocabulary, short-term memory, inhibition and CES-D), none of these dimensions seems to be involved when considering the role of the emotional pattern on verbal false memories.

Although we would need to further explore the role of emotionality on verbal false memory for semantically related words, we can sum up our results inferring that older adults don't seem to differ from younger adults concerning the role played both by arousal and by valence, here unprecedently analyzed in interaction.

References.

Brainerd, C., Stein, L., Silveira, R., Rohenkohl, G., & Reyna, V. (2008). How Does Negative Emotion Cause False Memories*? Psychological Science, 19* (9), 919–925. https://doi.org/10.1111/j.1467-9280.2008.02177.x

Brainerd, C. J., Yang, Y., Toglia, M. P., Reyna, V. F., & Stahl, C. (2008). Emotion and false memory: The Cornell/Cortland norms. In *annual meeting of the Psychonomic Society, Chicago, IL.*

Burgess, P. W., & Shallice, T. (1996). Response suppression, initiation and strategy use following frontal lobe lesions. *Neuropsychologia, 34* (4), 263–272. https://doi.org/10.1016/0028-3932(95)00104-2

Deese, J. (1959) On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology, 58*,17–22. DOI: 10.1037/h0046671

Dehon, H., Larøi, F., & van der Linden, M. (2010). Affective valence influences participant's susceptibility to false memories and illusory recollection. *Emotion, 10* (5), 627–639. https://doi.org/10.1037/a0019595

Devitt, A. L., Tippett, L., Schacter, D. L., & Addis, D. R. (2016). Autobiographical memory conjunction errors in younger and older adults: Evidence for a role of inhibitory ability. *Psychology and Aging, 31*(8), 927–942. https://doi.org/10.1037/pag0000129

Fernandes, M., Ross, M., Wiegand, M., & Schryer, E. (2008). Are the memories of older adults positively biased? *Psychology and Aging, 23*(2), 297–306. https://doi.org/10.1037/0882-7974.23.2.297

Jacoby, L. L., & Rhodes, M. G. (2006). False Remembering in the Aged. Current Directions in Psychological Science, 15(2), 49–53. https://doi.org/10.1111/j.0963-7214.2006.00405.x

Kensinger, E. A., & Corkin, S. (2004b). The effects of emotional content and aging on false memories. *Cognitive, Affective, & Behavioral Neuroscience, 4* (1), 1–9. https://doi.org/10.3758/cabn.4.1.1

Kunzmann, U., Kupperbusch, C. S., & Levenson, R. W. (2005). Behavioral Inhibition and Amplification During Emotional Arousal: A Comparison of Two Age Groups. *Psychology and Aging, 20* (1), 144–158. https://doi.org/10.1037/0882-7974.20.1.144

Lange K, Kühn S, Filevich E (2015) Correction: "Just Another Tool for Online Studies" (JATOS): An Easy Solution for Setup and Management of Web Servers Supporting Online Studies. *PLOS ONE 10* (7): e0134073. https://doi.org/10.1371/journal.pone.0134073

Lecce, S., Zocchi, S., Pagnin, A., Palladino, P., & Taumoepeau, M. (2010). Reading Minds: The Relation Between Children's Mental State Knowledge and Their Metaknowledge About Reading. *Child Development, 81*(6), 1876–1893. https://doi.org/10.1111/j.1467-8624.2010.01516.x

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. doi:10.3758/s13428-011-0168-7

L. Milfont, T., & Fischer, R. (2010). Testing measurement invariance across groups: applications in cross-cultural research. *International Journal of Psychological Research, 3* (1), 111–130. https://doi.org/10.21500/20112084.857

Norman, K. A., & Schacter, D. L. (1997). False recognition in younger and older adults: Exploring the characteristics of illusory memories. *Memory & Cognition, 25* (6), 838–848. https://doi.org/10.3758/BF03211328

Otani, H., Libkuman, T. M., Widner, R. L., & Graves, E. I. (2007). Memory for Emotionally Arousing Stimuli: A Comparison of Younger and Older Adults. *The Journal of General Psychology, 134* (1), 23–42. https://doi.org/10.3200/genp.134.1.23-42

Piguet, O., Connally, E., Krendl, A. C., Huot, J. R., & Corkin, S. (2008). False memory in aging: Effects of emotional valence on word recognition accuracy. *Psychology and Aging, 23* (2), 307–314. https://doi.org/10.1037/0882-7974.23.2.307

Radloff, L. S. (1977). The CES-D scale: A self report depression scale for research in the general population. *Applied Psychological Measurements,* 1, 385-401.

RStudio Team (2020). RStudio: Integrated Development for R. RStudio, PBC, Boston, MA URL http://www.rstudio.com/

Reed, A. E., & Carstensen, L. L. (2012). The Theory Behind the Age-Related Positivity Effect. *Frontiers in Psychology, 3*. https://doi.org/10.3389/fpsyg.2012.00339

Reyna, V. F., Corbin, J. C., Weldon, R. B., & Brainerd, C. J. (2016). How fuzzy-trace theory predicts true and false memories for words, sentences, and narratives. *Journal of Applied Research in Memory and Cognition, 5* (1), 1–9. https://doi.org/10.1016/j.jarmac.2015.12.003

Roediger, H. L., III McDermott, K. B. (1995). Creating false memories: Remembering words not presented in list. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 803–814. DOI: 10.1037/0278-7393.21.4.803

Roediger, H. L., Balota, D. A., & Watson, J. M. (2001). Spreading activation and arousal of false memories. In R.G. Crowder and H.L. Roediger (Eds.), The nature of

remembering: Essays in honor of Robert G. Crowder (pp. 95-115). Washington DC: American Psychological Association.

Sharkawy, J. E., Groth, K., Vetter, C., Beraldi, A., & Fast, K. (2008). False Memories of Emotional and Neutral Words. *Behavioural Neurology, 19* (1–2), 7–11. https://doi.org/10.1155/2008/587239

Soylu, F. (2010). Forward/backward digit-span task. Arch. Neurobehav. Exp. Stimuli 222. http://www.neurobs.com/ex_files/expt_view?id=218

Sutherland, M. R., & Mather, M. (2015). Negative arousal increases the effects of stimulus salience in older adults. *Experimental Aging Research, 41* (3), 259–271. https://doi.org/10.1080/0361073X.2015.1021644

Zhang, W., Gross, J., & Hayne, H. (2016) The effect of mood on false memory for emotional DRM word lists. *Cognition and Emotion 31*, 526–537. DOI: 10.1080/02699931.2016.1138930

CHAPTER 7

General Discussion.

Summary of results

The overaching aim of the present thesis was to disentangle the effects of emotionality on semantics in bilinguals and older adults through the study of verbal false memories.

The doctoral thesis began by describing verbal false memories, usually assessed through the DRM paradigm (Deese, 1959; Roediger & McDermott, 1995), as a means by which to measure verbal false memories, namely the effect according to which participants are likely to remember also words that are not part of a list, in light of their semantic association with the actually encoded materials. It has been argued that this kind of verbal false memories arises as a consequence of semantic processing, as expressed by theories like Fuzzy Trace Theory (FTT; Reyna *et al.*, 2016) and Spreading Activation Theory (SAT; Roediger *et al.*, 2001). Testing the semantic properties of the DRM paradigm, Brainerd and colleagues (2008) found that semantic activation spreading can reliably explain the occurrence of verbal false memories.

When the presented words have an emotional quality to them (e.g., shark, happy), this acts as a booster for semantic processing (Zhang, 2016). This effect has been addressed both when considering arousal and valence. In fact, the higher emotionality conveyed by the material elicits a higher number of verbal false memories (Brainerd *et al.*, 2010), so as for the negative valence words in comparison with the positive valence ones (e.g., Dehon, Larøi & Van der Linden, 2010).

The effects of both arousal and valence on false memories have been tested on young monolinguals (Sharkawy, Groth, Vetter, Beraldi, & Fast, 2008; Dehon, Larøi, &

Van der Linden, 2010). However, to the best of our knowledge, there is no literature on the effects of arousal and valence on the production of verbal false memories' by bilinguals, while at the same time their effects on aging have so far resulted in only few, inconsistent data.

Given this lack of research, we focused on the study of the populations of bilingual and older adults for the purpose of the present thesis. Interestingly enough, these two populations display opposite characteristics concerning the two main domains involved in semantic processing, language and memory.

As Bialystok demonstrated in her 2009 study, relative to monolinguals, the cognitive processing of bilinguals may benefit from advantages in executive functioning (including working memory, as in Baddeley & Hitch model, 1974), while simultaneously suffering in terms of vocabulary breadth (i.e., a bilingual may know a smaller number of words in each of the languages spoken, in comparison to monolingual peers).

In contrast, in healthy aging, especially episodic memory and executive functioning are deteriorating, while vocabulary knowledge seems to be preserved and, in some cases enhanced (e.g., Birren, 2006; Ramscar, Hendrix, Shaoul, Milin, & Bayeen, 2014).

We assessed the role of word emotionality in verbal false memory production and its relation to semantic processing on the connotative dimension (Allan, 2007) in these two populations by administering to all participants an Italian version of the Cornell/Cortland Emotional Lists /Brainerd *et al.*, 2008) that had been validated through a pilot study. All participants did a recognition test using these lists, which included the

actually coded words, their critical lures, and some distractors, divided in weakly semantically related and in semantically unrelated.

The associated empirical research of this thesis is organized in two main sections. Section 1 regarded bilinguals, section 2 regarded older adults. In section 1, we tested two main bilingual populations using the Italian version of the Cornell/Cortland emotional lists. One participant group was composed of children aged from 9 to 11 years and spoke a minority language as L1 and Italian as L2. The other group consisted of German/Italian bilingual adults living in Bozen, a fully bilingual city located in the North of Italy. For both the groups the DRM lists were in their L2, namely Italian. We chose to include these two different populations because they mainly differ on a dimension that turned out to be determinant when speaking about being bilinguals, namely socioeconomic status (SES) (Bonifacci *et al.*, 2020).

In chapter 2, we described a study on minority language bilingual children. The study's main results can be summed up as follows. Minority language bilingual children did neither differ from their monolingual peers in terms of verbal false memories, nor in terms of effect of emotionality on verbal false memory production. Coherent with the literature (e.g., Sharkawy, Groth, Vetter, Beraldi, & Fast, 2008; Dehon, Larøi, & Van der Linden, 2010), high arousal and negative valence lures were more falsely recognized in both groups than low arousal and positive valence lures. In contrast, the factor SES turned out to be significant in explaining the differences in verbal false memory. The results allow the conclusion that the type of bilingualism considered does not affect the occurrence of verbal false memories, neither in itself, nor in its emotional pattern. Minority language bilingual children do not seem to differ from their monolingual peers in terms of semantic processing, not even in its deep emotional component.

Because we found SES to be a significant factor, we ran another study described in chapter 3 to clarify the role of SES in minority language bilingual children relative to monolingual peers, aiming to assess its relation with vocabulary skills and reading comprehension skills. Both of these are determinants for assessing literacy skills. It turned out that SES was significant in explaining both vocabulary skills and reading comprehension skills. At the same time, when considering the relation between SES and reading comprehension, vocabulary size mediated this relation; and when considering the relation between SES and vocabulary skills, reading comprehension mediated that relation. From the combined results, we conclude that the three variables all interact in determining the development of literacy skills.

In chapter 4, we reported a study about German/Italian bilingual adults living in Bozen. Different from chapter 2, being bilingual or not now turned out did affect the occurrence of verbal false memories. In particular, bilinguals were less prone to produce verbal false memories than monolinguals. Concerning the emotional pattern, low arousal and positive valence were significant in reducing verbal false memories in the whole sample. However, this result was more marked for the bilingual sample. In fact, our results proved that the L2 of this bilingual population develops with less emotionality than in monolingual peers. This is an indication that this group of bilinguals develop semantic representations in a more 'shallow' way with respect to their L2.

As a whole, the findings reported in this section clarify that one should not generalize patterns about bilingual functioning on the basis of a single variable or factor. Clearly, many dimensions contribute to being bilingual. In particular, SES, the status of the spoken languages, and how they are recognized in the social context, all determine the

effects of being bilingual act on cognitive processing and, in this specific case, on semantics.

In section 2, we studied how a group of older adults responded to the Italian version of the Cornell/Cortland emotional lists in comparison with younger adults.

In chapter 5, we ran a major study aiming to resolve an open debate in the literature (e.g., Devitt & Schacter, 2016; Jacoby & Rhodes, 2006; Norman & Schacter, 1997). Why do older adults produce an increased number of verbal false memories? Is this because they rely more on semantic processing (e.g., Tun *et al.*,) or is because of the impairment in the cognitive and mnestic domain that is typical of aging (e.g., Askey & Playfoot, 2018)? To answer these questions, we assessed the semantic similarity of items included in the DRM paradigm. We also added cognitive control measures (i.e., inhibition skills and working memory) in order to assess their influence on false memory production in both the test group of older adults and the control group of younger adults.

At first, our results confirmed the precondition according to which older adults might produce more verbal false memories than younger adults. Then, we considered the role played by semantic similarity of the items on verbal false memories. It turned out that the more semantically associated words are the more falsely recognized ones in both younger and older adults. This result pattern was in line with the literature stating that DRM false memories arise in light of semantic processing (Brainerd *et al.*, 2008).

Concerning the role on semantics with respect to the age group, despite a trend towards significance of the interaction between semantic similarity and older group, the pattern seems to be comparable between younger and older adults. Even if these data seem not to allow us to consider semantics as the only responsible for the increased

number of verbal false memories in older adults, further insights come from the results concerning our inhibition measure. Indeed, inhibition skills were significant only for the older adults, but the direction of the effect was the contrary to what we expected: Better inhibition skills were associated with a higher number of verbal false memories in aging. More in detail, older adults who performed better in the inhibition task were the ones who produced a larger number of verbal false memories. We ascribed this finding to the task we used (the Hayling Test, Burgess & Shallice, 1996), which can also be intended as a measure of semantic control. These results seem to indicate that semantic activation in aging cannot be intended only as a compensatory mechanism for compromised memory skills and executive functioning. Indeed, is seems to be a resource they can rely on net of a functioning executive functioning domain.

In chapter 6, we considered the emotional component of the lexicon in the DRM paradigm and its effect on older adults' verbal false memories. With the caveat that effects of emotionality were not significant, we would still like to briefly discuss the study's results. We in particular wish to consider the trend towards significance that we found in the two-way interaction of low arousal and positive valence of the critical lures, associated to a reduction of verbal false memories in both the groups. When examining the direction of the effects of arousal and valence, we found no differences concerning the effect of emotionality between the two groups. Higher arousal lures were equally more prone to be falsely remembered than lower arousal ones both by younger and older adults, while negative valence did affect both groups in producing more false memories. These results are consistent with the literature on emotional false memories in younger adults (e.g., Sharkawy, Groth, Vetter, Beraldi, & Fast, 2008; Dehon, Larøi, & Van der Linden, 2010; Brainerd, Stein, Silveira, Rohenkohol, & Reyna, 2008). However, the results are inconsistent with the few literature considering arousal and valence as detached aspects of

processing (e.g., Kensinger & Corkin, 2004; Piguet, Connally, Krendl, Huot, & Corkin, 2008; Fernandes, Ross, Wiegand, & Schryer, 2008).

Future directions

In future research pertaining bilinguals, at first we would like to expose children of the same age as tested in chapter 2 (those of 9-11 years) to a semantic training based on the Italian version of the emotional DRM lists (CEL lists, Brainerd et al., 2008). We already formulated some stories using the items of the DRM emotional lists in a narrative.

Consequently and Following list design, the stories should be divided according to their arousal and their valence properties (i.e., high arousal and negative valence stories, high arousal and positive valence stories, low arousal and negative valence stories and low arousal a positive valence stories). In particular, we would like to expose children to a pre-training phase and a post-training phase. Both would consist of the same experimental design as in chapter 2, namely the emotional version of the DRM paradigm, with added control measures (i.e., vocabulary task and categorical fluency task, direct and backward span tasks). The aim would be to measure the verbal false memory pattern and the emotional pattern before and after the training. The hypothesis is that, the more the meanings of the words are learned with richness and complexity, the more a participant would be able to remember these words with precision. Also, as a side effect, the whole context in which the term is inserted would be activated, leading to the remembrance of words that were never actually presented. Given that the false memory effect for DRM lists results from increased semantic processing, we hypothesize that a semantic memory training that favors contextualization of presented terms, will lead to a larger production of false memories.

Another future direction concerning bilinguals would be to extend the study described in chapter 4 to a population of German/Italian bilingual children living in Bozen, comparable in terms of age to the population we tested in chapter 2. We would expect that the observed result pattern in this case would be slightly different. Growing up, people preferably choose one of their two languages as their main one. As a consequence many people in Bozen become unbalanced bilinguals as adults, preferring one of their two languages as more often used. In contrast, we hypothesize that bilingual children living in Bozen would be more balanced bilinguals. Therefore, studying this population would allow us a better comparison to the minority language bilingual children already investigated. The study would also provide important new insights into the development of balanced bilinguals.

For older adults, there could be follow-ups of the reported two studies. With respect to the study described in chapter 5, we could replicate our experimental design including another measure of inhibition skills, detached from semantics (i.e., the Stroop task). This would result in a better disentanglement of the effects of semantics and executive functioning. With respect to the study described in chapter 6, it might be useful to run the study again on a broader sample of participants, to increased our power to detect the effect of emotionality on verbal false memories in aging.

Concerning the role of emotion across all the studies described above, for the purposes of the present thesis we chose to embrace the role of emotionality as a means by which to better understand semantic processing. Indeed, we decided to consider the emotionality conveyed by the words included in our materials as an average of the answers that emerged from our pilot studies concerning arousal and valence attributions. We did not include other materials aimed at deepening the subjective emotional traits of

our participants, in order to avoid burdening the proposed task in terms of time. However, we cannot exclude a role played by the individual differences that could have affected the way in which the words are experienced. It might be of interest for future studies to also consider this aspect and to include it in our analysis. Indeed, we think it might be another new line to be followed in the literature concerning verbal false memory domain. As far as we reviewed, the literature so far investigate effects that are mainly oriented on how emotionality acts on verbal false memories being part of a context in which participants are immersed (e.g., Storbeck & Clore, 2011) or, on the contrary, on how emotionality acts on verbal false memories as intrinsical feature of the words participants have to remember (see Bookbinder & Brainerd, 2016 and Zhang, 2016 for reviews).

Conclusions

The general aim of the present thesis was to better understand the connotative dimension of semantics, in particular, the role that the emotionality conveyed by the meaning of words has on their processing. We studied this issue by examining verbal false memories obtained through the Italian version of the Cornell/Cortland emotional DRM lists by Brainerd and colleagues (2008). Two populations were tested that have opposite cognitive and linguistic features, namely bilinguals and older adults. The former group is advantaged in terms of executive functioning and working memories processing, but disadvantaged in terms of vocabulary breadth knowledge (Bialystok, 2009). The latter group is characterized by the opposite pattern, namely having executive functioning impaired by aging and a vocabulary size being equal or even better in comparison with younger adults (Birren, 2006).

From section 1, we conclude that the impact of being bilingual on verbal false memories for emotional words, largely depends on the kind of bilinguals we are

considering. Indeed, socioeconomic status of the person and the linguistic status of the language spoken (namely, the relevance that the society in which bilinguals live attributes to the languages they speak) seem to determine how semantic processing changes in comparison with monolingual peers.

From section 2, on the impact of aging on verbal false memories, we argue, first, that older adults produce more verbal false memories, because of a larger reliance on semantic processing at the expense of reduced abilities in mnestic and executive functioning domains. Secondly, based on our preliminary results, we argue that emotionality seems to affect verbal false memory in a similar way as in younger adults.

Verbal false memory has been extensively studied in the literature on various populations, not only because it is proof of our fallible memory, but mainly as a resource to better understand its embedding in semantics. The role of the emotionality conveyed by the words on this semantic processing is still poorly analyzed. In particular, there is no or a limited literature concerning the two main populations involved in the present thesis, namely bilinguals and older adults. To close this gap in the literature, we ran the studies discussed in the chapter of the thesis, for two mean reasons. First, we chose the two populations given the specific characteristics of each. In fact, for both groups there still is no unanimity concerning their semantic skills and, in particular, concerning the connotative dimension of these semantic skills. Furthermore, contrasting the results from these two radically different populations can provide insights and be a basis for future developments concerning semantics per se.

About bilingual people, we can conclude that they develop their semantic skills similarly to monolinguals. Indeed, we can conclude that semantics develops as partially detached from language skills in bilinguals. Nevertheless, it is not possible to generalize

this result to every bilingual, since the way in which people experiment their languages largely depends on the social context in which they are immersed.

About older adults, we can conclude that semantics is not affected by aging. Indeed, it seems to remain intact, both in itself and in its connotative dimension. Nevertheless, even if older adults rely more on semantic processing as a compensatory mechanism for their impaired memory and executive functioning domains, the latter still play a significant role in affecting their performance.

Semantics lies on the border between language and memory. However, it seems to develop as detached from both, as we assessed studying it in bilinguals and in older adults. It is complex and multifaced. Indeed, it deserves further explorations. We think that studying other populations that are peculiar in terms of their cognitive functioning can help isolating semantics and better understanding its edges.

References

Allan, K. (2007). The pragmatics of connotation. *Journal of pragmatics, 39* (6), 1047-1057.

Askey, C., & Playfoot, D. (2018). Examining theories of cognitive ageing using the false memory paradigm. *Quarterly Journal of Experimental Psychology, 71* (4), 931–939. https://doi.org/10.1080/17470218.2017.1307433

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. H. Bower (Ed.),The psychology of learning and motivation (Vol. 8, pp. 47–90). London: Academic Press.

Bialystok, E. (2009). Bilingualism: The good, the bad, and the indifferent. *Bilingualism: Language and cognition, 12* (1), 3-11.

Birren, J. E. Handbook of the Psychology of Aging, Sixth Edition (Handbooks of Aging) (2005–12-23). (2021). Academic Press; 6 edition (2005–12-23).

Bonifacci, P., Lombardo, G., Pedrinazzi, J., Terracina, F., & Palladino, P. (2020). Literacy skills in bilinguals and monolinguals with different SES. *Reading & Writing Quarterly, 36* (3), 243-259.

Bookbinder, S. H., & Brainerd, C. J. (2016). Emotion and false memory: The context–content paradox. *Psychological Bulletin, 142(12)*, 1315–1351. https://doi.org/10.1037/bul0000077

Brainerd, C., Stein, L., Silveira, R., Rohenkohl, G., & Reyna, V. (2008). How Does Negative Emotion Cause False Memories? *Psychological Science, 19* (9), 919–925. https://doi.org/10.1111/j.1467-9280.2008.02177.x

Brainerd, C. J., Yang, Y., Toglia, M. P., Reyna, V. F., & Stahl, C. (2008). Emotion and false memory: The Cornell/Cortland norms. In *annual meeting of the Psychonomic Society, Chicago, IL.*

Brainerd, C., Holliday, R., Reyna, V., Yang, Y., & Toglia, M. (2010). Developmental reversals in false memory: Effects of emotional valence and arousal. *Journal of Experimental Child Psychology,* 107 (2), 137–154. https://doi.org/10.1016/j.jecp.2010.04.013

Burgess, P. W., & Shallice, T. (1996). Response suppression, initiation and strategy use following frontal lobe lesions. *Neuropsychologia, 34* (4), 263–272. https://doi.org/10.1016/0028-3932(95)00104-2

Deese, J. (1959) On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology, 58*,17–22. DOI: 10.1037/h0046671

Dehon, H., Larøi, F., & van der Linden, M. (2010). Affective valence influences participant's susceptibility to false memories and illusory recollection. *Emotion, 10* (5), 627–639. https://doi.org/10.1037/a0019595

Fernandes, M., Ross, M., Wiegand, M., & Schryer, E. (2008). Are the memories of older adults positively biased? Psychology and Aging, 23(2), 297–306. https://doi.org/10.1037/0882-7974.23.2.297

Jacoby, L. L., & Rhodes, M. G. (2006). False Remembering in the Aged. *Current Directions in Psychological Science, 15* (2), 49–53. https://doi.org/10.1111/j.0963-7214.2006.00405.x

Kensinger, E. A., & Corkin, S. (2004b). The effects of emotional content and aging on false memories. *Cognitive, Affective, & Behavioral Neuroscience, 4* (1), 1–9. https://doi.org/10.3758/cabn.4.1.1

Norman, K. A., & Schacter, D. L. (1997). False recognition in younger and older adults: Exploring the characteristics of illusory memories. *Memory & Cognition, 25*(6), 838–848. https://doi.org/10.3758/BF03211328

Piguet, O., Connally, E., Krendl, A. C., Huot, J. R., & Corkin, S. (2008). False memory in aging: Effects of emotional valence on word recognition accuracy. *Psychology and Aging, 23* (2), 307–314. https://doi.org/10.1037/0882-7974.23.2.307

Ramscar, M., Hendrix, P., Shaoul, C., Milin, P., & Baayen, H. (2014). The Myth of Cognitive Decline: Non-Linear Dynamics of Lifelong Learning. *Topics in Cognitive Science*, *6*(1), 5–42. https://doi.org/10.1111/tops.12078

Reyna, V. F., Corbin, J. C., Weldon, R. B., & Brainerd, C. J. (2016). How fuzzy-trace theory predicts true and false memories for words, sentences, and narratives. *Journal of Applied Research in Memory and Cognition, 5* (1), 1–9. https://doi.org/10.1016/j.jarmac.2015.12.003

Roediger, H. L., Balota, D. A., & Watson, J. M. (2001). Spreading activation and arousal of false memories. In R.G. Crowder and H.L. Roediger (Eds.), *The nature of remembering: Essays in honor of Robert G. Crowder* (pp. 95-115). Washington DC: American Psychological Association.

Roediger, H. L., III McDermott, K. B. (1995). Creating false memories: Remembering words not presented in list. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 803–814. DOI: 10.1037/0278-7393.21.4.803

Sharkawy, J. E., Groth, K., Vetter, C., Beraldi, A., & Fast, K. (2008). False Memories of Emotional and Neutral Words. *Behavioural Neurology, 19* (1–2), 7–11. https://doi.org/10.1155/2008/587239

Storbeck, J., & Clore, G. L. (2011). Affect influences false memories at encoding: Evidence from recognition data. *Emotion, 11*(4), 981–989. https://doi.org/10.1037/a0022754

Tun, P. A., Wingfield, A., Rosen, M. J., & Blanchard, L. (1998). Response latencies for false memories: Gist-based processes in normal aging. *Psychology and Aging, 13* (2), 230–241. https://doi.org/10.1037/0882-7974.13.2.230

Zhang, W., Gross, J., & Hayne, H. (2016) The effect of mood on false memory for emotional DRM word lists. *Cognition and Emotion 31*, 526–537. DOI: 10.1080/02699931.2016.1138930