This is a preprint of the manuscript entitled "I remember emotional content better, but I'm struggling to remember who said it!", which has been accepted for publication in Cognition. Minor differences between this version and the final version of the article might be found. For the final version, please refer to the actual publication:

Le Bigot, L., Knutsen, D., & Gil, S. (2018). I remember emotional content better, but I'm struggling to remember who said it! *Cognition, 180,* 52-58. doi:10.1016/j.cognition.2018.07.001

I remember emotional content better, but I'm struggling to remember who said it!

Ludovic LE BIGOT*[1], Dominique KNUTSEN[2], and Sandrine GIL[1]

[1]Université de Poitiers, France ; CNRS (CeRCA UMR 7295), France

[2]Department of Psychology, University of Essex, UK

*Correspondence about this article should be addressed to:

Ludovic Le Bigot

Centre de Recherches sur la Cognition et l'Apprentissage (CeRCA), Bât. A5, MSHS, 5 rue Théodore Lefebvre, TSA 21103, 86073 Poitiers Cedex 9, France.

Tel.: +33 5 49 45 46 22

Fax: +33 5 49 45 46 16

E-mail: ludovic.le.bigot@univ-poitiers.fr

Abstract

The joint impact of emotion and production on conversational memory was examined in two experiments where pairs of participants took turns producing verbal information. They were instructed to produce out loud sentences based on either neutral or emotional (Experiment 1: negative; Experiment 2: positive) words. Each participant was then asked to recall as many words as possible (content memory) and to indicate who had produced each word (reality monitoring). The analyses showed that both self-production and emotion boost content memory, although emotion also impairs reality monitoring. This study sheds light on how both factors (emotion and production) may constrain language interaction memory through information saliency.

Keywords: conversational memory; emotion; production effect; dyadic interaction; reality monitoring

I remember emotional content better, but I'm struggling to remember who said it!

1. Introduction

You and your colleague are talking about a dinner organized at your boss's house tomorrow. As the interaction unfolds, you should both encode information about what was said, although there is evidence that you and your partner might subsequently remember this information differently. Researchers have typically addressed this issue by investigating unemotional conversations, even though emotion colors our daily life experience. Indeed, imagine that no one likes your boss and that negative information has been exchanged about this dinner, or imagine that your boss is great and that someone mentions that they are very excited about tomorrow night. Are you more likely to remember negative, positive or neutral information? An additional question concerns memory for who said what. Indeed, you might remember a piece of information well, but could you accurately say whether you produced this information yourself, or whether it was produced by someone else, depending on its emotional valence? This study seeks to examine how memory processes and emotional content jointly impact content memory and memory for who said what in a conversation-like setting.

1.1. The impact of emotion and production on memory for conversational content

Memory plays a central role in human conversation. Indeed, the contributions produced during any interaction are usually encoded in each participant's memory. This information may then be resorted to during subsequent interactions to support dialogic partner-adaptation (for examples, see Clark & Schaefer, 1989; Clark & Wilkes-Gibbs, 1986; Horton & Brennan, 2016; Horton & Gerrig, 2005, 2016; Kronmüller & Barr, 2015). Various factors may affect memory for the content of an interaction, including the nature of the partners' relationship (acquaintances vs. friends; Samp & Humphreys, 2007), or whether they share the same job status (e.g., Holtgraves, Srull, & Socall, 1989). In this context, the fact that this previous work has seldom investigated the link between conversational memory and emotion is surprising, as some authors have already pointed out that emotion could be the key to understanding conversational memory. For instance, Keenan, MacWhiney, and Mayhew (1977) wrote that "findings that interactional content improves memory can be explained [by] the affective nature of high interactional content statements" (p. 558-559). Why study conversational memory as an unemotional construct, when it inherently results from social interaction – thus necessarily involving both emotion and cognition (see Keltner & Horberg, 2015)?

Key to the proposal that emotion can influence conversational memory is the evidence that emotional words are memorized better than neutral ones in standard memory tasks involving free recall (e.g., Talmi & Moscovitch, 2004), short-term memory (e.g., Monnier & Syssau, 2008) or recognition (e.g., Thapar & Rouder, 2009). The few studies which have directly examined the influence of emotion on conversational memory per se have shown, for instance, dialogue partners recall conversational content more accurately after pleasant interactions (Samp & Humphreys, 2007).

It is also important to point out at this stage that one of the key features of any conversation is that both (or more) conversational partners have the opportunity to produce utterances during the interaction. This involves that from each partner's point of view, some utterances are self-produced whereas others are partner-produced. This has a major impact on conversational memory, due to a *production* effect in memory. This term refers to the fact that information produced out loud is remembered better than information read silently or produced by someone else (MacLeod, 2011; MacLeod, Gopie, Hourihan, Neary, & Ozubko, 2010). This effect has been generalized to spontaneous dyadic interactions (Knutsen & Le

Bigot, 2014; 2015; McKinley, Brown-Schmidt, & Benjamin, 2017; Yoon, Benjamin, & Brown-Schmidt, 2016; see also Knutsen & Le Bigot, 2017), revealing that each conversational partner tends to remember what he or she said better than what the other person said after the end of the interaction. However, once again, one limitation of this work is that it has focused solely on unemotional conversations (map tasks, matching tasks involving Tangram figures, etc.), making it difficult to determine whether self-production affects conversational memory regardless of emotion.

1.2.The impact of emotion and production on conversational reality monitoring

During any conversation, partners memorize not only *what was said*, but also *who said what*. In a basic, one on one conversation, this implies being able to distinguish between internally versus externally generated utterances (e.g., Fischer, Shult, & Steffens, 2015; Johnson & Raye, 1981; Raye & Johnson, 1980). This ability, which is not specific to conversation, is usually referred to as *reality monitoring*. The results of related past research on the effect of emotion on memory for contextual information such as reality monitoring are somewhat mixed. Some researchers have reported that although emotional content is memorized better than neutral content, contrasted patterns are found when participants are asked to remember who said what (i.e., a disadvantage, or no particular effect, for emotional items; e.g., Davidson, McFarland, & Glisky, 2006); in contrast, other studies have found that emotion causes participants to remember better who said what (e.g., D'Argembeau & Van der Linden, 2004; Doerksen & Shimamura, 2001). However, the information provider was not systematically an actual person in these studies, limiting the generalization of the findings to conversational settings. What is more, none of these studies involved situations in which the participant provided some of the information him- or herself. In Davidson et al.'s (2006) study, the information was provided by one of two prerecorded voices (one male, one female). In D'Argambeau and Van der Linden (2004), and Doerksen and Shimamura's (2001) studies, the "source" of the information was operationalized as a feature of the target word (i.e., ink or background color). These limitations imply that the effect of emotional valence on conversational reality monitoring has not yet been examined directly.

Reality monitoring has also been examined in the context of research on the production effect. When a piece of information benefits from self-production, reality monitoring is less efficient, as self-production causes the identity of the provider of the information to be remembered less well (e.g., Fischer et al., 2015; Jurica & Shimamura, 1999; although see also McKinley et al., 2017, who reported no significant effect of production on reality monitoring). The contrast between content memory and reality monitoring is in line with a content-context trade-off hypothesis, whereby concomitant encoding of content (e.g., what was said) and contextual information (e.g., who said what) causes competition for limited cognitive resources (Jurica & Shimamura, 1999; Nieznanski, 2011). However, once again, the information provider was not systematically an actual person in previous studies, nor did the participant generate information him- or herself, limiting the generalization of the findings to conversational settings. For instance, in Jurica and Shimamura's study, participants interacted with faces shown on a computer screen, rather than actual people.

1.3.The current study

The current study sought to overcome the limitations of previous related work by examining the combined effect of emotion and production on participants' memory for words (emotional vs. neutral) produced either by themselves or by another participant in a conversation-like setting. The participants' performance on a subsequent reality monitoring task was also examined. Although the participants did not have the opportunity to engage in spontaneous conversation (which would have made the emotional content of their utterances difficult to control), they did have the opportunity to take turns producing information. This study also sought to examine whether the joint effect of emotion and production on memory is found for negative (Experiment 1) and positive (Experiment 2) content. The latter point was addressed for two reasons. Firstly, consistent with the negativity-bias literature, some studies suggest that negative information is more likely to be processed automatically and to have an influence on psychological functioning as a whole (for a review, see Baumeister, Bratslavsky, Finkenauer, & Vohs, 2001). Secondly, studies have highlighted that although both negative and positive stimuli are more likely to be remembered than neutral ones, negative emotions make stimuli details particularly salient, at the expense of contextual information (Kensinger, 2009). This is consistent with research on the *weapon focus effect* (for a review, see Fawcett, Russell, Peace, & Christie, 2013), whereby the presence of a weapon on an event decreases memory for peripheral information. Consequently, the nature of what is remembered may vary depending on the valence of the information stored in memory, suggesting a potential modulation of the content-context trade-off hypothesis. The predictions were that self-produced and emotional words are better recalled than partner-produced and neutral words; the opposite pattern should be found for reality monitoring.

2. Experiment 1

2.1. Method

2.1.1. Rationale

In Experiment 1, pairs of participants were first informed that they would have to perform a collaborative task together. They were then shown neutral and negative words and took turns to produce out loud sentences which included these words. After this, each partner was asked to recall as many of these words as possible (content memory) and to indicate who had produced each word (or reality monitoring). Finally, the interaction ended with the collaborative task which the partners had previously been told about and during which they elaborated a short story together based on the information they had memorized. The data from this final phase were not of prime interest here; thus, they were not analyzed (the sole purpose of this phase was to emphasize the collaborative dimension of the experiment to the participants). The main analysis sought to examine the influence of production (i.e., whether the words initially shown on screen were self- or partner-produced) and emotion (i.e., whether the words initially shown on-screen were negative or neutral) on the participants' performance on the recall and reality monitoring tasks.

2.1.2. Participants

Forty-six University students (42 women; *Mean age* = 20.33, *SD* = 1.73) provided informed consent before taking part in the study in exchange for course credit or payment and were divided into 23 dyads. Four participants were removed from the final sample because they did not follow the instructions or they were not native French speakers, thus resulting in a sample of 42 participants in 21 dyads (38 women; *Mean age* = 20.86, *SD* = 2.39).

2.1.3. Materials and procedure

112 nouns were selected from the Affective Norms for French Words (Monnier & Syssau, 2014), which provides emotional valence, arousal, imageability and book and film frequencies. Examples of the nouns used in Experiment 1 (and Experiment 2) are provided in Table 1. Eight neutral nouns were used as examples and 104 nouns (52 negative and 52 neutral) were used in the remainder of the study. The two categories of words differed in terms of emotional valence and arousal, which were measured on a 9-point scale; the mean valence ratings were 2.28 (*SD* = .41) for the negative nouns and 4.93 (*SD* = 1.49) for the

neutral nouns. Mean arousal ratings were 5.20 (*SD* = 2.76) for the negative nouns and 3.46 (*SD* = 2.26) for the neutral nouns. The negative and neutral nouns did not significantly differ in length (i.e., the same proportion of 1-, 2- and 3 syllable-words was the same), in imageability (*M* = 4.67, *SD* = 1.40; *M* = 4.60, *SD* = 1.05, respectively), in book frequency (*M* = 55.92, *SD* = 79.29; *M* = 53.22, *SD* = 90.78, respectively) or in film frequency (*M* = 55.04, *SD* = 92.75; *M* = 40.32, *SD* = 65.84, respectively), all *Fs* < 1 (ANOVA).

Table 1

Example of Words used in Experiments 1 and 2

Example of negative words (Experiment 1)	Example of positive words (Experiment 2)	Example of neutral words (Experiments 1 and 2)
Haine (hatred)	Cadeau (gift)	Théorie (theory)
Douleur (pain)	Rose (rose)	Artichaut (artichoke)
Araignée (spider)	Miracle (miracle)	Code (code)
Poubelle (bin)	Chocolat (chocolate)	Cintre (hanger)
Lassitude (weariness)	Energie (energy)	Gauche (left)
Traître (traitor)	Ambition (ambition)	Echarpe (scarf)
Squelette (skeleton)	Tendresse (kindness)	Machine (machine)
Ambulance (ambulance)	Liberté (freedom)	Transport (transport)

Note. English translations are provided in brackets.

Pairs of participants sat in a quiet room and were informed that they would be “partners” for the remainder of the study (i.e., one participant was referred to as “Partner 1” and the other as “Partner 2”); they were also told that the study would end with a collaborative task. Each partner sat in front of a computer screen on either side of a partition so that they could hear but not see each other (this was to control for the use of nonverbal cues such as head nods, hand gestures, etc.). Moreover, during the entire experimental session, the participants were recorded using two microphones connected to a digital voice recorder. The experimenter stayed in the room to manage the different phases.

The experimental design was divided into five phases, two of which were of particular interest in this study: the collaborative production phase (phase 1) and the conversational memory assessment phase (phase 3) (both content memory and reality monitoring were assessed during the latter phase). Phases 2 and 4 were interfering tasks (respectively, the Symbol search and the Coding subtests of the Wechsler Adult Intelligence Scale; WAIS-IV; Wechsler, 2008). Both interfering tasks took two minutes to complete; the resulting data were not analyzed. Phase 5 consisted in a collaborative writing task.

During the collaborative production phase, the participants used nouns which were presented to them on the computer screens. The program used to present the words was written in E-Prime 2.1. The two screens were connected, which meant that the two partners were shown the same stimuli at the same time.

In each trial, a fixation cross was presented on the screens for 1000ms; it was then followed by the presentation of a noun for 1500ms. At this point, the partners also received information (i.e., “Partner 1” or “Partner 2”) about whose turn it was to produce a sentence out loud, using the noun presented on the screen. After the partner had produced the sentence, the experimenter used a wireless mouse to move on to the next trial. Each participant thus produced 52 sentences out loud (26 were based on a negative noun and 26 were based on a neutral noun). The nouns and the information about whose turn it was to produce the sentence were presented in a random order.

Following the collaborative production phase, both partners independently completed the first interfering task for two minutes (the experimenter used a stopwatch to keep track of the time). The partners then embarked on the conversational memory assessment phase, during which each participant was asked to write on a sheet of paper as many words as possible which had been presented on screen during the collaborative production phase (content memory). The experimenter interrupted them after five minutes. The participants then had to indicate whether the words they remembered had been self- or partner-produced (reality monitoring) by writing “self” or “partner” next to each recalled word. The participants were not allowed to communicate during this phase. The participants then completed the second interfering task before embarking on the final collaborative task.

2.1.4. Data coding and experimental design

The participants’ memory for target words (i.e., words which had been presented on-screen and produced during the collaborative production phase) and for who had produced them was assessed by examining their performance during the conversational memory assessment phase. Each target word was coded either as recalled (code 1) or non-recalled (code 0). This was a binary variable. The words recalled correctly were then coded for reality monitoring: the participant’s response (self- or partner-produced) was coded either as correct (code 1) or incorrect (code 0). This was also a binary variable.

The two main independent variables (IVs) were Valence (negative, neutral) and Production (from each participant’s point of view: self-produced, partner-produced). Both IVs were within-participants.

Moreover, recall that the participants’ task during the collaborative production phase was to produce sentences which included the target words. The length of the sentences produced by the participants could have affected their subsequent memory for these words. In order to discard this possibility, all sentences produced during the collaborative production phase were transcribed and coded for content words (see Table 2 for an example). This category included common names (e.g., “cat”), proper names (e.g., “Paris”), adjectives (e.g., “small”) and verbs (e.g., “eat”). 300 sentences (representing 13.74% of the entire dataset, and selected randomly) were double-coded for content words; the coders reached an initial agreement level of 92% *(Cronbach’s Alpha* = 0.98). All disagreements were discussed and resolved, and the remainder of the data was single-coded. The number of content words initially produced per sentence was used as a continuous, centered IV in this study, in order to ensure that the effects of Valence and Production remained unchanged even when this additional variable was taken into account.

Table 2

Content Word Sample Coding

Target word	Valence	Sentence produced	Content words
Ligne (line)	Neutral	suivre la ligne n'est pas important (following the line is not very important)	suivre, ligne, être, important (following, line, be, important)
Larme (tear)	Negative	une larme coule sur ta joue (a tear rolls on your cheek)	larme, couler, joue (tear, roll, cheek)

Note. English translations are provided in brackets.

2.2. Results

The analysis was conducted in SAS 9.4 (GLIMMIX procedure). Logistic mixed models were used to analyze the data. Logistic models are used to analyze data from experiments in which the outcome variable is binary, which was the case here (i.e., target nouns were either recalled or not, and reality monitoring responses were either correct or not). In this case, these models were used to calculate odd ratios, which quantify the probability of an event (e.g., correctly recognizing a target noun) occurring relative to another event (e.g., failing to recognize a target noun).

As for mixed models, they include random intercepts, which account for potential variability across dyads, participants and items (i.e., nouns), and random slopes, which account for the fact that dyads, participants and items may differ in their sensitivity to within-unit IVs (by-dyad random effects were included in this study because the participants completed the collaborative production phase in pairs; see McMahon, Pouget, & Tortu, 2006). The maximal random structure justified by the experimental design (i.e., all random intercepts and all random slopes corresponding to within-unit IVs) was initially implemented, in line with Barr, Levy, Scheepers, and Tily's (2013) recommendations. Then, the random effects causing G-matrix convergence failure were identified and removed from the model (the identification of problematic random effects is performed automatically in SAS). Removing these random effects from the model has no effect on the outcome of the analysis (i.e., even if the degrees of freedom of the model are higher when these random effects are removed, the parameters of the model remain unchanged; Kiernan, Tao, & Gibbs, 2012). The results reported hereafter correspond to the final model.

All main effects were systematically included in the analyses; interactions were only included if they reached statistical significance. Finally, there were six cases in which the participants did not produce a sentence during the collaborative production phase, due to an experimenter error. These occurrences were discarded from further analysis. Because of this, the number of observations in each cell of the design varied slightly across cells; a Satterthwaite correction was applied to the degrees of freedom in order to account for this. Moreover, the second analysis (reality monitoring) was only performed on a subset of the data, as highlighted below. The correction was also applied to account for this in this analysis.

Content memory. The mean proportion of words correctly recalled during the conversational memory assessment phase is reported in Figure 1 (left panel). The final model used to analyze the data included Valence, Production and the Number of content words as fixed effects. The random effects structure included by-participant and by-item random intercepts and by-dyad random slopes corresponding to the Number of content words.

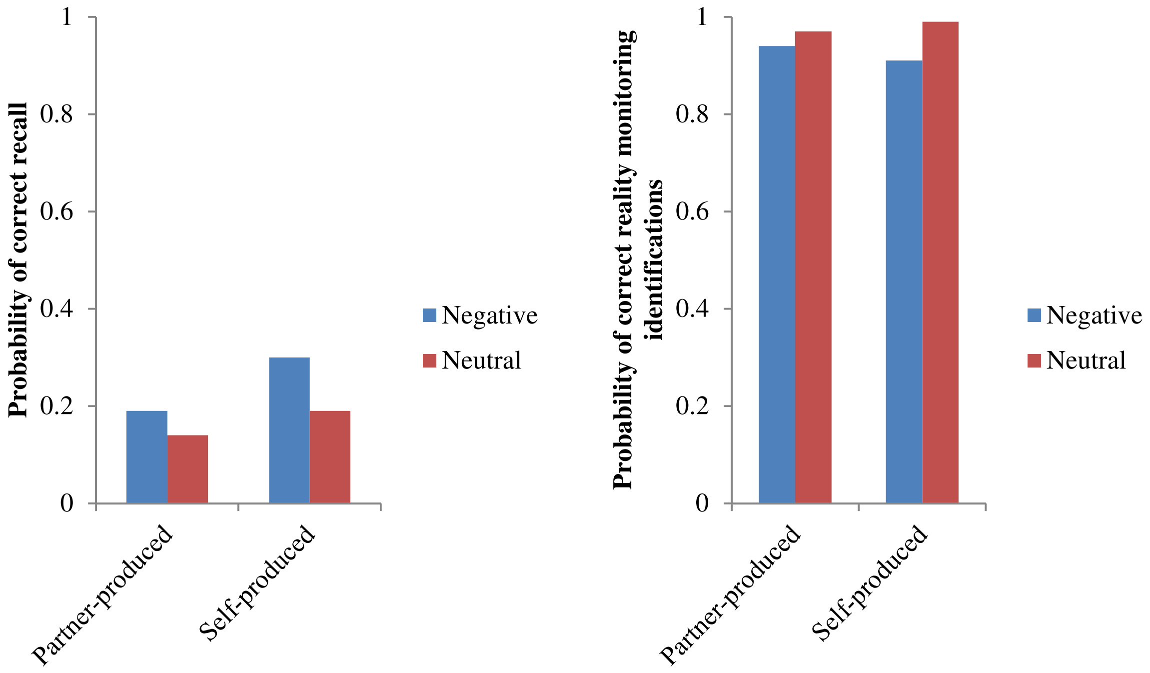

Figure 1. Proportion of correct recall (left panel) proportion of correct reality monitoring identifications (right panel) as a function of Presentation and Valence in Experiment 1.

Valence and Presentation significantly predicted content memory, $F(1, 104) = 11.60$, $p < .001$, and $F(1, 4358) = 44.18$, $p < .001$. The participants were more likely to recall negative words than neutral words, $OR = 1.67$, $CI_{.95} = 1.24, 2.26$. They were also more likely to recall self-produced words than partner-produced words, $OR = 1.70$, $CI_{.95} = 1.45, 1.98$. The effect of the Number of content words failed to reach statistical significance, $F(1, 24) = 3.19$, $p = .087$.

Reality monitoring. The mean proportion of correct reality monitoring identifications during the conversational memory assessment phase is reported in Figure 1 (right panel). This represents a conditional probability – only the nouns which had been recalled previously were included in this second analysis. The final model used to analyze the data included Valence, Production and the Number of content words as fixed effects. The random effects structure included only by-item random intercepts and by-participant random slopes corresponding to the Valence.

Valence significantly predicted reality monitoring, $F(1, 569) = 10.70$, $p = .001$. Participants were less likely to provide a correct response when the word was negative than when it was neutral, $OR = 0.19$, $CI_{.95} = 0.07, 0.52$. The effects of Production and Number of content words failed to reach statistical significance, respectively $F(1, 886) = 0.17$, $p = .681$ and $F(1, 25) = 0.99$, $p = .330$.

In sum, the results of Experiment 1 partly supported the hypotheses. As predicted, self-produced information was better recalled than partner-produced information; negative information was also better recalled than neutral information. Emotion affected reality monitoring, as the participants were less likely to remember who had said what when the word produced was negative than when it was neutral. No conclusion can be drawn regarding the potential effect of production on reality monitoring, as no significant effect was found.

The purpose of Experiment 2 was to attempt to replicate these findings by examining the participants' memory for *positive* and neutral words, extending the generalizability of the results to different kinds of emotions (i.e., negative vs. positive valence).

3. Experiment 2

3.1. Method

3.1.1. Participants

Fifty-four University students (53 women; *Mean age* = 21.03, *SD* = 6.40) participated in exchange for course credit or payment and were divided into 27 dyads. Eight participants were removed because participants were not native French speakers or they had not followed the instructions, thus resulting in a sample of 46 participants in 23 dyads (all female; *Mean age* = 21.28, *SD* = 6.91).

3.1.2. Materials, procedure and experimental design

The procedure and experimental design were the same as those used in Experiment 1, except that negative nouns were replaced by positive nouns. As before, 52 positive nouns were selected from the Monnier and Syssau (2014) dataset. Examples of the nouns used in Experiment 2 are provided in Table 1. The neutral nouns used were the same as Experiment 1. The two categories of nouns (positive vs. neutral) did not differ in length (i.e., there was the same proportion of 1-, 2- and 3 syllable-words). Positive nouns did not differ from neutral ones in imageability (M = 5.05, SD = 1.58), in book frequency (M = 74.66, SD = 71.89) or in film frequency (M = 63.50, SD = 69.59), all $Fs < 1$ (ANOVA). As in Experiment 1, the positive nouns category differed significantly from the neutral category in terms of emotional valence – with positive nouns being more positive (M = 7.89, SD = .47) – and in terms of arousal – with positive nouns having a higher degree of arousal (M = 6.36, SD = .78).

3.2. Results

The data were analyzed following the same rationale as in Experiment 1. Twenty-two cases in which the participants did not produce a sentence during the collaborative production phase, due to an experimenter error, were discarded from further analysis.

Content memory. The mean proportion of words correctly recalled during the conversational memory assessment phase is reported in Figure 2 (left panel). The final model used to analyze the data included Valence, Production and the Number of content words as fixed effects. The random effects structure included by-dyad, by-participant and by-item random intercepts, by-dyad random slopes corresponding to Valence, by-participant and by-item random slopes corresponding to Production and by-item random slopes corresponding to the Number of content words.

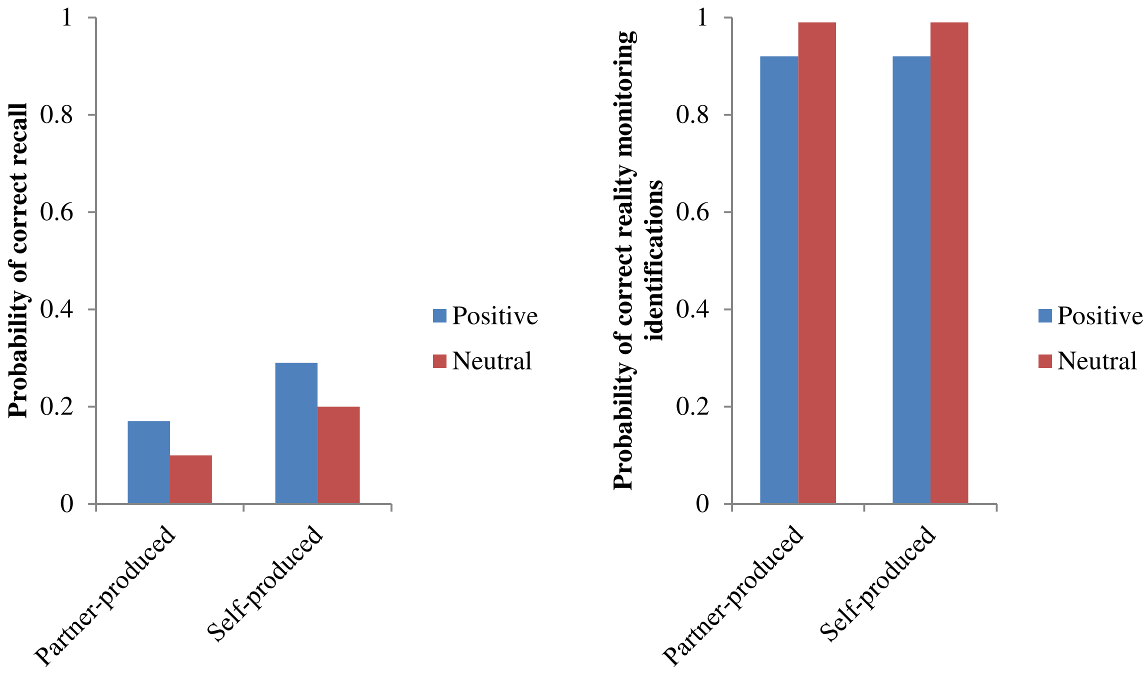

Figure 2. Proportion of correct recall (left panel) proportion of correct reality monitoring identifications (right panel) as a function of Presentation and Valence in Experiment 2.

Valence and Presentation significantly predicted content memory, $F(1, 63) = 15.95$, $p < .001$, and $F(1, 40) = 60.72$, $p < .001$. The participants were more likely to recall positive words than neutral words, $OR = 1.73$, $CI_{.95} = 1.32, 2.28$. They were also more likely to recall self-produced words than partner-produced words, $OR = 2.21$, $CI_{.95} = 1.80, 2.71$. This pattern of results replicates the findings of Experiment 1 in a situation where positive and neutral words were compared.

The Number of content words also significantly predicted content memory, $F(1, 102) = 23.47$, $p < .001$. The probability of correct recall increased as the number of content words initially produced increased, $b = .27$.

Reality monitoring. The mean proportion of correct reality monitoring identifications during the conversational memory assessment phase is reported in Figure 2 (right panel). The final model used to analyze the data included Valence, Production and the Number of content words as fixed effects. The random effects structure included by-dyad, by-participant and by-item random intercepts, and by-dyad random slopes corresponding to Production.

Valence significantly predicted reality monitoring, $F(1, 886) = 14.38$, $p < .001$. Participants were less likely to provide a correct response when the word was positive than when it was neutral, $OR = 0.13$, $CI_{.95} = 0.05, 0.37$. The effects of Production and Number of content words failed to reach statistical significance, respectively $F(1, 20) < .001$, $p = .911$ and $F(1, 886) = 0.82$, $p = .366$. This pattern of results replicates that found in Experiment 1, extending it to positive versus neutral items.

4. Discussion

The purpose of the current study was to examine the combined impact of emotion and production on memory in a conversation-like setting in which two participants took turns producing utterances.

Firstly, the results extend the finding that emotional information is remembered better than neutral information (for a review, see Hamann, 2001) to language interaction contexts. This effect could be due to emotion impacting the encoding of stimuli through the involvement of specific attention and perception mechanisms (Easterbrook, 1959), leading to subsequent enhanced memory performance. Indeed, the emotional properties – which are processed early and automatically – of any stimulus capture attention (for a review, see Vuilleumier & Driver, 2007). Moreover, in contrast to a valence-based asymmetrical point of view, which can be summarized as “bad is stronger than good” (Baumeister et al., 2001), emotional content (either positive or negative) was recalled better than neutral content in both experiments.

From a theoretical perspective, the results of both experiments suggest that the effect of emotion on language interaction memory can be explained both in terms of valence (e.g., Kensinger, 2009) or arousal (e.g., Mather & Sutherland, 2009), as emotional nouns were more arousing than neutral ones in the current study. There is a large debate as to the most influential dimension of emotion (i.e., valence or arousal) on memory (for a discussion, see Kensinger, 2004). Generally speaking, emotional stimuli are also often more arousing than neutral ones, making it difficult to disentangle the respective contribution of these two emotional dimensions to psychological functioning. However, a small number of studies have revealed that emotional words are memorized better than neutral ones even in the absence of differences in arousal (Adelman & Estes, 2013). Moreover, the only study to have directly examined the effect of arousal – as measured by electrodermal reactivity – on conversational memory concluded that there is no reliable impact of the former on the latter (MacWhinney, Keenan, & Reinke, 1982).

Secondly, the present work sheds light on the dissociation between content memory and memory for who said what. Importantly, this study was one of the first to examine the joint influence of emotion and production on reality monitoring in a situation in which the information providers were actual people who took turns producing information. The main finding here was that the information provider (self or other) was more likely to be identified correctly for neutral words than for valenced ones. This offers a better understanding of the somewhat mixed findings reported in the literature (e.g., Davidson et al., 2006; Doerksen & Shimamura, 2001) – as suggested in the Introduction section, this discrepancy might be due to “sources” being operationalized differently across studies. It is also important to note that most of these studies did not focus on reality monitoring per se, as they often required participants to distinguish between two external sources of information. We suggest that the detrimental effect of emotion on reality monitoring found in this study is consistent with the well-documented finding that emotion enhances memory for information deemed central, but has no influence or affects negatively memory for peripheral information (for an exhaustive review, see Levine & Edelstein, 2010). Consistently with Easterbrook's (1959) attention hypothesis, both types of information (i.e., what was said vs. who said what) compete for cognitive resources at the time of encoding. When a piece of information is emotionally charged, it attracts attention, thus becoming central from the speaker's point of view. In this situation, few or no resources are left to process who produced this information (such contextual information would be deemed peripheral in this situation), explaining why the source was identified less well for emotional words in this study.

Thirdly, the results for content memory suggest that distinct and complementary cognitive mechanisms underlie both the production and the emotion effects in conversation-like settings; these two mechanisms would then work together towards increasing information accessibility in conversational memory. As mentioned previously, the current study was undertaken in order to offer a better understanding of the psychological processes at play during dialogue. Indeed, although the experimental setup used did not allow the participants to engage in genuine, spontaneous verbal interactions, this study nonetheless examined the participants' content memory and reality monitoring in a context where the information memorized was produced by actual people (i.e., oneself or the other participant) taking turns producing utterances. In this context, the basic memory biases found in this study are likely to affect the management of genuine interactions (Horton & Gerrig, 2005), as the level of accessibility of information stored in conversational memory plays a central role in subsequent partner-adaptation (Clark & Marshall, 1981; Horton & Gerrig, 2005).

With that in mind, as a first step, this research replicates previous findings on the production effect in conversation-like contexts, as participants remembered self-produced words better than partner-produced words (see MacLeod, 2010). This is in line with other experiments on dialogue, which have suggested that the effect of self-production on content memory can cause two conversational partners to hold quite different memories of what was said during past interactions (Fischer et al., 2015; Knutsen & Le Bigot, 2012, 2014, 2017; Yoon et al., 2016). Then, the main and most original contribution of this work is that it shows that emotionally valenced information is likely to be particularly salient in the partners' common ground, making them potentially more likely to reuse this information for adaptation purposes in subsequent interactions; however, this is at the expense of reality monitoring. Future research focusing on situations which are more similar to real-life communication will provide a better understanding of the joint effects of emotion and production (and the underlying mechanisms) in dialogue.

The lack of a significant effect of production on reality monitoring in the current study must be interpreted with caution. Indeed, the proportion of correct reality monitoring identifications was very high in both experiments, potentially yielding a ceiling effect which would have made any production effect harder to detect. Even more importantly, had a significant effect of production been detected, this would have been difficult to interpret in the current study. Indeed, a response bias towards saying "partner" more often than "self" in the reality monitoring task would have yielded the same pattern of results as the expected reversed production effect (i.e., the "partner" response would have been more likely to be produced than the "self" response). This kind of issue could have been solved if a recognition task (i.e., involving new items – words which were not actually presented to the participants in the first place) had been used to assess content memory instead of a recall task. Thus, at this point, no firm conclusion can be drawn regarding the effect of production on reality monitoring. In any event, the current findings support the idea that the two components of conversational memory (i.e., content memory and reality monitoring) must be considered separately when studying the impact of the features of the interaction situation. Indeed, we found that although both components are affected by emotion, this effect was in opposite directions in the two analyses.

Other limitations to the current study also open avenues for future research. In particular, the question of which aspect(s) of emotion (valence or arousal) affect(s) memory will be addressed in more detail. Besides, the samples used in this study included a majority of female participants (who thus formed same-gender dyads). Previous work suggests that gender may be considered as a salient feature in an interaction, leading to poorer reality monitoring in same-gender than in mixed-gender dyads (e.g., Fischer et al., 2015; Macrae, Bodenhausen, & Calvini, 1999). This social aspect of the interaction will also be taken into account in future studies.

In conclusion, this study was guided by the idea that human conversational memory is constrained by information saliency. The goal of the study was to determine whether such saliency depends both on who produced the information in the first place, and on its emotional content. The use of more ecological communication situations in future research will highlight how the social context in which the interaction takes place moderates these findings. In any event, this work suggests that even if might you remember well that someone predicted that the dinner at your boss's house would be a complete disaster, or a fantastic success, you might have difficulty remembering who mentioned this emotional information.

Acknowledgements

The authors thank Marie-Françoise Crété, Inès Fergant and Ioana Sandulescu for their assistance in data collection.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Bibliography

Agresti, A. (2002). *Categorical data analysis* (2nd ed). New York, NY: John Wiley & Sons.

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

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology, 5,* 323.

Clark, H. H. (1996). *Using language.* Cambridge, MA: Cambridge University Press.

Clark, H. H., & Marshall, C. (1981). Definite reference and mutual knowledge. In A. K. Joshi, B. L. Webber, & I. A. Sag (Eds.), *Elements of discourse understanding* (Vol. 2, pp. 10–63). Cambridge, MA: Cambridge University Press.

Clark, H. H., & Schaefer, E. F. (1989). Contributing to discourse. *Cognitive Science*, *13*, 259-294.

Clark, H. H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition, 22*, 1–39.

D'Argembeau, A., & Van der Linden, M. (2004). Influence of affective meaning on memory for contextual information. *Emotion*, *4*, 173-188.

Davidson, P. S., McFarland, C. P., & Glisky, E. L. (2006). Effects of emotion on item and source memory in young and older adults. *Cognitive, Affective, & Behavioral Neuroscience*, *6*, 306-322.

Doerksen, S., & Shimamura, A. P. (2001). Source memory enhancement for emotional words. *Emotion*, *1*, 5-11.

Easterbrook, J. A. (1959). The effect of emotion on cue utilization and the organization of behavior. *Psychological Review*, *66*, 183-201.

Fawcett, J. M., Russell, E. J., Peace, K. A., & Christie, J. (2013). Of guns and geese: A meta-analytic review of the ‘weapon focus’ literature. *Psychology, Crime & Law*, *19*, 35-66.

Fischer, N. M., Schult, J. C., & Steffens, M. C. (2015). Source and destination memory in face-to-face interaction: A multinomial modeling approach. *Journal of Experimental Psychology: Applied*, *21*, 195-204.

Hamann, S. (2001). Cognitive and neural mechanisms of emotional memory. *Trends in Cognitive Sciences*, *5*, 394-400.

Holtgraves, T., Srull, T. K., & Socall, D. (1989). Conversation memory: The effects of speaker status on memory for the assertiveness of conversation remarks. *Journal of Personality and Social Psychology*, *56*, 149-160.

Horton, W. S., & Brennan, S. E. (2016). The role of metarepresentation in the production and resolution of referring expressions. *Frontiers in Psychology, 7.*

Horton, W. S., & Gerrig, R. J. (2005). Conversational common ground and memory processes in language production. *Discourse Processes, 40*, 1–35.

Horton, W. S., & Gerrig, R. J. (2016). Revisiting the memory-based processing approach to common ground. *Topics in Cognitive Science, 8,* 780–795.

Horton, W. S., & Keysar, B. (1996). When do speakers take into account common ground? *Cognition, 59*, 91–117.

Johnson, M. K., & Raye, C. L. (1981). Reality monitoring. *Psychological Review*, *88*, 67-85.

Jurica, P. J., & Shimamura, A. P. (1999). Monitoring item and source information: Evidence for a negative generation effect in source memory. *Memory & Cognition*, *27*, 648-656.

Keenan, J. M., MacWhinney, B., & Mayhew, D. (1977). Pragmatics in memory: A study of natural conversation. *Journal of Verbal Learning and Verbal Behavior*, *16*, 549-560.

Keltner, D., & Horberg, E. J. (2015). Emotion-cognition interactions. In M. Mikulincer, P. R. Shaver, E. Borgida & J. A. Bargh (Eds.), *APA Handbook of Personality and Social Psychology* (Vol. 1, pp. 623-664). Washington, DC: American Psychological Association.

Kensinger, E. A. (2004). Remembering emotional experiences: The contribution of valence and arousal. *Reviews in the Neurosciences*, *15*, 241-252.

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

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

Kiernan, K., Tao, J., & Gibbs, P. (2012). Tips and strategies for mixed modelling with SAS/STAT procedures. Paper presented at the *2012 SAS Global Forum*, Orlando, FL.

Knutsen, D., & Le Bigot, L. (2012). Managing dialogue: How information availability affects collaborative reference production. *Journal of Memory and Language*, *67*, 326-341.

Knutsen, D., & Le Bigot, L. (2014). Capturing egocentric biases in reference reuse during collaborative dialogue. *Psychonomic Bulletin & Review*, *21*, 1590-1599.

Knutsen, D., & Le Bigot, L. (2015). The influence of reference acceptance and reuse on conversational memory traces. *Journal of Experimental Psychology: Learning, Memory and Cognition, 41,* 574-585.

Knutsen, D., & Le Bigot, L. (2017). Conceptual match as a determinant of reference reuse in dialogue. *Journal of Experimental Psychology: Learning, Memory and Cognition, 43,* 350-368.

Kronmüller, E., & Barr, D. J. (2015). Referential precedents in spoken language comprehension: A review and meta-analysis. *Journal of Memory and Language, 83,* 1-19.

Levine, L.J., & Edelstein, R.S. (2009). Emotion and memory narrowing: A review and goal relevance approach. *Cognition & Emotion*, *23*, 833–875.

MacLeod, C. M. (2011). I said, you said: The production effect gets personal. *Psychonomic Bulletin & Review, 18*, 1197–1202.

MacLeod, C. M., Gopie, N., Hourihan, K. L., Neary, K. R., & Ozubko, J. D. (2010). The production effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*, 671–685.

Macrae, C. N., Bodenhausen, G. V., & Calvini, G. (1999). Contexts of cryptomnesia: May the source be with you. *Social Cognition*, *17*, 273-297.

MacWhinney, B., Keenan, J. M., & Reinke, P. (1982). The role of arousal in memory for conversation. *Memory & Cognition*, *10*, 308-317.

Mather, M., & Sutherland, M. (2009). Disentangling the effects of arousal and valence on memory for intrinsic details. *Emotion Review*, *1*, 118-119.McKinley, G. L., Brown-Schmidt, S., & Benjamin, A. S. (2017). Memory for conversation and the development of common ground. *Memory & Cognition, 45*, 1281-1294.

McMahon, J. M., Pouget, E. R., & Tortu, S. (2006). A guide for multilevel modeling of dyadic data with binary outcomes using SAS PROC NLMIXED. *Computational Statistics & Data Analysis*, *50*, 3663-3680.

Monnier, C., & Syssau, A. (2008). Semantic contribution to verbal short-term memory: Are pleasant words easier to remember than neutral words in serial recall and serial recognition? *Memory & Cognition*, *36*, 35-42.

Monnier, C., & Syssau, A. (2014). Affective norms for French words (FAN). *Behavior Research Methods*, *46*, 1128-1137.

Nieznański, M. (2011). Generation difficulty and memory for source. *The Quarterly Journal of Experimental Psychology*, *64*, 1593-1608.

Raye, C. L., & Johnson, M. K. (1980). Reality monitoring vs. discriminating between external sources of memories. *Bulletin of the Psychonomic Society*, *15*, 405-408.

Rozin, P., & Royzman, E. B. (2001). Negativity bias, negativity dominance, and contagion. *Personality and Social Psychology Review*, *5*, 296-320.

Samp, J. A., & Humphreys, L. R. (2007). "I said what?" Partner familiarity, resistance, and the accuracy of conversational recall. *Communication Monographs*, *74*, 561-581.

Talmi, D., & Moscovitch, M. (2004). Can semantic relatedness explain the enhancement of memory for emotional words? *Memory & Cognition*, *32*, 742-751.

Thapar, A., & Rouder, J. N. (2009). Aging and recognition memory for emotional words: A bias account. *Psychonomic Bulletin & Review*, *16*, 699-704.

Vuilleumier, P., & Driver, J. (2007). Modulation of visual processing by attention and emotion: Windows on causal interactions between human brain regions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *362*, 837-855.

Wechsler, D. (2008). *Wechsler Adult Intelligence Scale*– Fourth Edition. San Antonio, TX: Pearson.

Yoon, S. O., Benjamin, A. S., & Brown-Schmidt, S. (2016). The historical context in conversation: Lexical differentiation and memory for the discourse history. *Cognition, 154,* 102–117.