Source and Destination Memory in Face-to-Face Interaction: 
A Multinomial Modeling Approach

Nele M. Fischer
Friedrich Schiller University Jena

Janette C. Schult
Friedrich Schiller University Jena

Melanie C. Steffens
University of Koblenz-Landau

Author Notes

Nele Fischer is now at the Department of Social Issues and Health, Leipzig University of Applied Sciences, Leipzig, Germany; Janette Schult, Institute of Psychology, Friedrich Schiller University Jena, Jena, Germany; Melanie Steffens, Department of Social and Economic Psychology, Faculty of Psychology, University of Koblenz-Landau, Landau, Germany.

This research was funded by a grant from the German Research Foundation to the senior author (DFG Ste 938/7-2). We thank Matthias Keil for programming, Parzival Eisner for help with data collection, Gabriele Engelhardt for language editing, and Sonja Gaukel, Nadine Knab, Colin MacLeod, Sandra Päpper, and David Riefer for valuable comments on a previous draft. Parts of this paper were presented at the TeaP, Vienna, 2012, and at the Meeting of the German Social Psychology Interest Group, Hagen, 2013. Please address correspondence to the senior author, Melanie Steffens, FB 8-Psychology, University of Koblenz-Landau, Fortstraße 7, 76 829 Landau/Pfalz, Germany. Phone: +49 6341 280 31214, Fax: +49 6341 280 31483, E-mail: steffens@uni-landau.de.
Abstract

Arguing that people are often in doubt concerning to whom they have presented which information, Gopie and MacLeod (2009) introduced a new memory component, destination memory: remembering the destination of output information (i.e., who did you tell this to?). They investigated source (i.e., who told you that?) versus destination memory in computer-based imagined interactions. The present paper investigates destination memory in real interaction situations. In two experiments in mixed-gender versus same-gender groups ($N=57$ and $N=89$), source and destination memory were manipulated by creating a setup similar to speed dating. In dyads, participants completed phrase fragments with personal information, taking turns. At recognition, participants decided whether fragments were new or old, and if old, whether they were listened to or spoken, and which depicted person was the source or the destination of the information. A multinomial model was used for analyses. Source memory significantly exceeded destination memory, whereas information itself was better remembered in the destination than in the source condition. These findings corroborate the trade-off hypothesis: The context is better remembered in input than output events, but the information itself is better remembered in output than input events. We discuss implications of these findings for real-world conversation situations.

Keywords: source monitoring; multinomial modeling; conversation memory; destination memory; generation effect
Source and Destination Memory in Face-to-Face Interaction:
A Multinomial Modeling Approach

Imagine that you are at a conference and talk to many colleagues every day sharing your exciting new research. On the final day of the conference, you run into a colleague with whom you talked at the welcome session. You want to tell her about your latest experiment but unfortunately you do not remember whether you have already told her. You are sure she would remember if you had, so it appears embarrassing to ask. Can memory research address this discrepancy in memory for a conversation topic? The aim of the present research is to investigate whether memory for the source of information (“Who told you that?”) exceeds memory for the target or destination (“To whom did you tell it?”) in conversation situations. In two experiments, we used a setup similar to speed dating. Providing the theoretical context, we first review the source-monitoring framework, the memory consequences of active involvement in information output, and destination memory.

A fact told to us by someone can be regarded as content of a conversation. In contrast, the person who has told the story, the location, and other details can be considered the context. Memory for content (item memory) and context (source memory) can be dissociated (Johnson, Hashtroudi, & Lindsay, 1993). Source monitoring is the process of evaluating source characteristics, including decision processes about the origin of memories. According to the source monitoring framework, a source is defined by the “… variety of characteristics that, collectively, specify the conditions under which a memory is acquired” (Johnson et al., 1993, p. 3), including memories that originate in oneself, such as speaking. External source monitoring is differentiating which external context attributes are related to the information, for example, with whom one talked about a topic. The discrimination between internally and externally generated memories such as “what one said aloud from what one heard would be classified as reality monitoring” (Johnson et al., 1993, p. 4). The source monitoring
framework does not explicitly differentiate between the direction of information transfer, such as speaking versus listening. Consequently, external source monitoring may refer to a source as the provider of information (sender, speaker) as well as to its destination or target (receiver, listener: Jurica & Shimamura, 1999; Marsh & Hicks, 2002). A third option is that a source is unspecific contextual detail (e.g., color of a text, Meiser & Bröder, 2002).

Source judgments are affected by response bias such as systematic guessing in situations of uncertainty about the source of a memory (e.g., Johnson & Raye, 1981; Klauer & Wegener, 1998; Johnson et al., 1993). Various conditions and activities affect the accuracy of source monitoring decisions (e.g., Johnson et al., 1993; Mulligan, 2004; 2011), among them speaking or listening in a conversation. Active involvement in information output (e.g., enacting, speaking aloud, completing words or phrases) typically increases item memory (e.g. compared to reading, listening, observing others). Generation effects (Slamecka & Graf, 1978), production effects (e.g., MacLeod, 2011; MacLeod, Gopie, Hourihan, Neary, & Ozubko, 2010; Ozubko, Gopie & MacLeod, 2012), and enactment effects (e.g., Rosa & Gutchess, 2011; see Engelkamp, 1998) underline the impact of active involvement on (recognition) memory. It is debatable whether they can always be distinguished unambiguously (see the enactment task list by Cohen, 1981). In addition to item memory, active involvement often increases reality monitoring, facilitating the discrimination between self and other as the origin of information (e.g., Ozubko et al., 2012; Riefer et al., 2007; Rosa & Gutchess, 2011). However, active involvement appears to hinder external source monitoring (e.g., Jurica & Shimamura, 1999; Koriat, Ben-Zur, & Druch, 1991; Mulligan, 2004; 2011; Riefer et al., 2007). In sum, active involvement typically increased item memory and reality monitoring, but decreased external source monitoring.

The fact that active involvement has implications for recognition memory and source monitoring has been discussed in the context of communication research (e.g., Brown, Jones,
Various studies investigated communication processes by means of the monitoring processes of a questioner and a responder or a provider and receiver of information (e.g., Brown et al., 1995; Jurica & Shimamura, 1999; Marsh & Hicks, 2002). Investigating source monitoring in conversations, participants were asked to monitor whether a certain statement was made by them or someone else (reality monitoring, Raye & Johnson, 1980), or which information was provided by whom (external source monitoring, see also the “Who-Said-What?” paradigm, Taylor, Fiske, Etcoff, & Ruderman, 1978). As several studies showed, findings depend on features of the conversation situation. For example, both source monitoring and reality monitoring are more difficult in same-gender than mixed-gender groups (e.g., Klauer & Wegener, 1998; Macrae, Bodenhausen, & Calvini, 1999). Jurica and Shimamura (1999) simulated a conversation on the computer (see also Marsh & Hicks, 2002). Three faces alternated telling participants person-oriented, depersonalized, or trivia statements or asking them to respond to questions. Answering questions generally increased item memory but impaired external source monitoring (i.e., identifying the related face), compared to reading. Similarly, external source monitoring was more error-prone than reality monitoring (Raye & Johnson, 1980).

How can the memory consequences of active involvement be explained? According to Koriat et al. (1991), active involvement in information output impairs external source monitoring by means of reduced context integration due to concentrating on what one is doing. In contrast, information input increases context integration because people focus on the contextual features that accompany the information. Input information is assumed to be perceived as belonging to the context it emerged from, associating it more strongly with contextual details, resulting in better item-context associations (i.e., external source monitoring). Output information instead is given out to the external context. Thus, it is
perceived as belonging more to oneself and is therefore not as well-integrated into the context. Consequently, the context is better remembered in input than output events, but the information is better remembered in output than input events: the trade-off hypothesis (Jurica & Shimamura, 1999). A similar mechanism regarding context memory has been called the attention hypothesis: Attention toward internal processes reduces, but externally oriented attention improves context memory (Gopie & MacLeod, 2009).

These authors coined the term destination memory for the “processes involved in remembering the destination of information that people output” (Gopie & MacLeod, 2009, p. 1492), while referring to the process of remembering the provider of incoming information as source memory. They compared source and destination memory in a computer-based between-subjects design. In the destination condition, participants “told” items aloud to photographs, while in the source condition, items were “presented” by photographs and read silently. Results were in line with the prediction that destination memory is poorer than source memory. These findings are reminiscent of the negative generation effects in external source monitoring introduced above (e.g., Jurica & Shimamura, 1999; McDaniel & Bugg, 2008; Riefer et al., 2007).

Although Gopie and MacLeod (2009) attributed high social relevance to destination memory, given their operationalizations, it is unclear whether their findings generalize to applied contexts (also see Lindner, Drouin, Tanguay, Stamenova, & Davidson, 2014). They presented pictures of famous people (sources and destinations) separately from items. This methodology is different from a face-to-face interaction in several relevant ways. First, in interactions, there is no separate presentation of items and pictures; this may have reduced item-picture associations. Second, as compared to interactions, seeing each picture on an otherwise identical computer screen provides fewer differential context cues that are known to enhance memory (e.g., Eich, 1985). Both of these procedures may have impaired both
source and destination memory as compared to applied contexts. Third, additional processes take place in real-world interactions, for example, nonverbal cues of active listening (“minimal responses”, e.g., Kollock, Blumstein, & Schwartz, 1985), and subtle feedback following statements (e.g., Local, 2005). These conversational features should specifically strengthen destination as compared to source memory. Fourth, memory in get-acquainted conversations may differ from that in other situations because the focus is on learning about the partner (e.g., Stafford & Daly, 1984). Finally, active involvement often has different consequences for memory in within-subject than between-subjects designs (e.g., MacLeod et al., 2010; Steffens & Erdfelder, 1998). Within-subject designs mimic conversations better than between-subject designs. For these reasons, we compared source and destination memory in a more ecological setting.

**Experiments 1a and 1b**

The present experiments investigated the relation between source and destination memory in real face-to-face interactions. We adapted Brown and colleagues’ (1995) experimental design: Participants alternated in asking and responding. Each group of participants was divided into three dyads. Within each dyad, participants either spoke by completing a phrase fragment (destination condition) or listened to the completion of their dyadic partner (source condition), close to a real conversation (see Kutsen & Le Bigot, 2014; MacLeod, 2011; Macrae, et al., 1999; Perfect, Field, & Jones, 2009, for similar setups). Following Gopie and MacLeod (2009, Exp. 2), phrase fragments were used as items that participants orally completed with personal information. After exchanging several items, partners were re-combined into new dyads until all participants were paired once. Concurrently, items, exposure time, and item-partner constellation were standardized. In a subsequent memory test participants decided whether they recognized the item, whether it was spoken or listened to, and indicated the dyadic partner associated with the item. For
analyses, we used a multinomial processing tree model to disentangle memory and guessing processes (e.g., Meiser & Bröder, 2002).

In line with the trade-off hypothesis (and the findings by MacLeod, 2011), we expected that item memory is increased in speaking compared to listening (Hypothesis 1). Following the trade-off hypothesis and Gopie and MacLeod (2009), destination memory should be impaired compared to source memory (Hypothesis 2). As frequently found for the generation effect (e.g., Riefer et al., 2007), reality monitoring (i.e., remembering whether an item was spoken or listened to) should be increased for spoken compared to listened-to material (Hypothesis 3). In addition to the memory effects, we expected a systematic guessing bias (e.g., Johnson & Raye, 1981; Ozubko et al., 2012; Riefer et al., 2007): Items should be attributed externally (“listened to”) rather than internally (“spoken”) if participants were unsure about their actual origin (Hypothesis 4).

Experiments 1a and 1b were identical except that mixed-gender groups participated in Experiment 1a, and either only men or only women attended each session of Experiment 1b. In mixed-gender groups, person memory should be better than in same-gender groups because of higher distinctiveness of conversation partners (Hypothesis 5, e.g., Klauer & Wegener, 1998). Finally, reality monitoring should be superior when interacting in different-gender rather than same-gender dyads (Hypothesis 6, e.g., Macrae et al., 1999).

Method

Experiment 1a: Participants

At a large German university, 57 participants were recruited ($M_{age} = 23.39$, $SD = 4.89$, 77% women) in mixed-gender groups of six. In four of the 10 groups, there were two men, in another two groups, there were three men. The four single men who only interacted with women were excluded from analyses, leaving 53 participants.

Experiment 1b: Participants
One participant with poor German language skills was excluded, leaving a total of 89 participants ($M_{age} = 24.06$, $SD = 5.16$, 66% women). In both experiments, if fewer than six participants showed up, a male confederate helped out (3 times in Exp. 1a) whose data were excluded from analyses. Participants received €3 or course credit.

**Materials and Procedure**

Participants were invited to a study about “the impact of first impressions”. Initially, a standardized digital portrait photograph with neutral facial expression of each participant was taken for the recognition test. All computer-based programs were written in PsycScope X B57 (Cohen, MacWhinney, Flatt, & Provost, 1993) and presented on Apple iBooks (13 inch).

**Study Items.** Phrase fragments were used as study items (e.g., “My favorite movie is ...”; “My hometown is...”). The topics included individual (dis)likes, opinions, hobbies, and university issues. Sixty target items were used during study and an additional 60 distractors were used for recognition; both sets were comparable regarding length and complexity. To select appropriate items, a pre-test was conducted with $n=20$ additional participants who rated items’ complexity and appropriateness in conversations with strangers. We randomly assigned 60 items to the target and 60 to the distractor list, matched for topic, out of a pool of 232 phrase fragments (60 fragments used by Gopie & MacLeod, 2009, and 172 self-generated fragments). The target item list was divided into 10 subsets, each containing six items matched for length, complexity, and topic.

**Setting.** The room was divided into three labeled study cabins and a separate testing area. Each cabin was equipped with a bar table with the opposite sides labeled 1 and 2 at which complementary study booklets were placed. The session started with comprehensive instructions. Participants drew a number between one and six that defined which subsets were spoken and which were listened to with which conversational partner at which position. Participants were instructed not to talk, except for the demanded item completion; also, they
were asked to turn a page only when requested and not to move or close the study booklet. Then participants were grouped into three dyads, each facing each other at one of the tables.

**Study Phase.** Item presentation and procedure were organized by the study booklet. First, a summary of instructions was presented. The speaking task was to read aloud each item and to complete it immediately with personal information in a short, informative way (“Please complete: I would love to travel to …”). Meanwhile the conversational partner’s booklet showed the instruction to listen to the item. Each item was presented on a separate page in the booklet and alternated with a page with the listening instruction. Items were shuffled to avoid simultaneous processing within different dyads. After presenting one subset, each booklet instructed participants to change to the appropriate next position. The experiment started with a sound signal that indicated to turn one page of the booklet. Every 13 seconds, a sound signal asked all participants to turn a page. After each participant spoke and listened to six items, respectively, a sound signal instructed them to change positions. After 60 sec, a signal started the next item completion phase. This procedure was repeated until all participants partnered once with every other participant.

**Recognition Test.** The testing area contained six cubicles equipped with numbered computers for a surprise recognition test without time limit. Photographs had been imported into the testing program. The instruction was displayed as the starting screen. Then, all target and distractor items were presented in a random order. Individual item completions were omitted: Only the phrase fragments as printed in the book were presented (Gopie & MacLeod, 2009; Jurica & Shimamura, 1999; see Greenwald & Johnson, 1989, for a generation effect for such cues). Up to three different subtask screens were presented for each item. First, participants decided whether the item had been presented at study or not. The response “new” initiated the presentation of the next item. If participants responded “old”, they were asked whether the item was spoken or listened to. Finally, the five conversation
partner photographs were presented with the question “To whom did you tell the fact?” or “Who told you the fact?”. Then the next item was presented. After the recognition test, demographics were collected, and participants were thanked, rewarded and debriefed. The experiment took about 45 minutes.

**Design**

The independent variable was speaking versus listening (within subject). Dependent variables were item memory, reality monitoring (correct responses to previously spoken items as “spoken” and listened-to items as “listened to”), and partner identification (selection of the correct conversational partner).

**The Present Model**

Measuring recognition memory requires data correction methods because of false positive responses due to guessing (e.g., Bröder & Meiser, 2007). We used a multinomial processing tree model (Batchelder & Riefer, 1990; Riefer & Batchelder, 1988) suitable to disentangle recognition from guessing processes and item memory from various source-monitoring processes (see Bröder & Meiser, 2007). These models have been adapted to various research purposes in measuring source memory (Batchelder, Hu, & Riefer, 1994; Klauer & Wegener, 1998; Meiser & Bröder, 2002; for a review see Erdfelder, Auer, Hilbig, Aßfalg, Moshagen, & Nadarevic, 2009). The observed categorical response pattern is considered the result of several cognitive processes estimated by the model’s parameters. If a threshold is crossed, the item or source is correctly identified. Otherwise, the item or source remains unidentified. Observed responses for unidentified items or sources result from guessing. The present model has three trees (spoken, listened-to, and new items; see Figure 1). It is based on the crossed-source model (Meiser & Bröder, 2002), further specifications being derived from Riefer and colleagues (2007) as well as Klauer and Wegener (1998).
All significance tests were conducted with $p < .05$. To facilitate comparison with previous research, we first present analyses based on the general linear model.

**General Linear Model for Experiment 1a**

Following Mulligan (2004), in recognition corrected hit rates were computed as the proportion of hits minus the proportion of false alarms. For reality monitoring and partner identification, we computed the proportions of correct source or destination identifications for items that were previously correctly identified as old.

Regarding *item memory*, a 2 (speaking vs. listening) × 2 (male vs. female conversation partner) ANOVA showed that speaking led to significantly better memory than listening (estimated marginal means: $M_s = .97$ and .90, $SEs = .004$ and .010), $F(1,52) = 63.32, p < .001$, $\eta^2_p = .55$, corroborating Hypothesis 1 (both other $F$s < 1.15). The same 2×2 ANOVA on *partner identification* showed that identification was higher for listened-to than for spoken items ($M_s = .94$ and .82, $SEs = .009$ and .020), $F(1,52) = 49.35, p < .001$, $\eta^2_p = .49$, corroborating Hypothesis 2 that destination memory was poorer than source memory.

Additionally, a main effect of the conversation partner’s gender indicated that person identification was better for (the fewer) male compared to female conversation partners ($M_s = .91$ and .85, $SEs = .016$ and .017, $F(1,52) = 8.33, p < .01$, $\eta^2_p = .14$; interaction: $F < 1$: source memory: $M_s = .96$ and .92, destination memory: $M_s = .85$ and .79, respectively). *Reality monitoring* was close to perfect. A 2 (speaking vs. listening) × 2 (same-gender vs. mixed-gender dyad) ANOVA indicated no significant difference between listened-to and spoken items ($M_s = .99$, $SEs = .003$), nor did reality monitoring differ when interacting in same versus mixed gender dyads ($M_s = .99$, $SEs = .002$ and .004, respectively), nor was there an interaction (all $F$s < 2.36, $ps > .13$). Thus, Hypotheses 3 and 6 could not be corroborated.

**General Linear Model for Experiment 1b**
A one-way ANOVA on item memory showed that, as in Experiment 1a, spoken items were better remembered than listened-to items ($M_s = .95$ and $.87$, $SEs = .007$ and .013), $F(1,88) = 89.75$, $p < .001$, $\eta^2_p = .51$, confirming Hypothesis 1. Again confirming Hypothesis 2, partner identification for listened-to items was better than for spoken items ($M_s = .83$ and .70, $SEs = .021$ and .022), $F(1,88) = 89.17$, $p < .001$, $\eta^2_p = .50$, indicating that destination memory was worse than source memory. Partner identification both regarding destination and source memory was lower in same-gender groups (Exp. 1b) than in mixed-gender groups (Exp. 1a, $t[133] = 3.47$, $p < .001$, and $t[132] = 3.88$, $p < .001$, respectively), in line with Hypothesis 5. Regarding reality monitoring, the overall hit rate was slightly higher for identifying listened-to items as listened to compared to identifying spoken items as spoken ($M_s = .97$ and .96, $SEs = .011$ and .009). This difference missed statistical significance, $F(1,88) = 3.62$, $p = .06$, and was in the opposite direction predicted by Hypothesis 3.

**Multinomial Modeling Analyses**

**Parameter Specifications of Model**

Response frequencies in each of 13 categories were aggregated over participants (see Table 1). The parameter $DOS (DOL)$ in Figure 1 represents the probability of identifying an old spoken (listened to) item as old. If identified as old, the conditional probability of identifying a previously spoken item as spoken (listened-to item as listened to) (reality monitoring) is represented by parameter $dS (dL)$. If not identified as spoken (1-$dS$) or listened to (1-$dL$), the response was guessed. With probability $bL (1-bL)$, the source “listened to” (spoken) was guessed. The parameters $pS$ and $pL$ indicate whether the correct partner was identified given reality monitoring ($dL$ or $dS$). If reality monitoring failed and participants guessed whether an item was spoken or listened to ($bL; 1-bL$), they identified the correct partner with probability $eL$ or $eS$. The parameters $pS$ and $eS$ represent partner identification for spoken items (i.e., destination memory). The parameters $pL$ and $eL$ represent partner
identification for items listened to (i.e., source memory). If not identified (with probabilities $1-pS, 1-eS, 1-pL, \text{ or } 1-eL$), the correct partner was guessed with the probability $1/n$, where $n$ is the number of partners. If the item was not recognized ($1-DOS; 1-DOL$), it was guessed to be old ($bO$) or new ($1-bO$). If the item was guessed old, all further responses resulted from guessing. The trees for spoken and listened-to items are built alike, with respective labeling of parameters. Guessing processes ($bO, bL, 1/n$) are identical across conditions, representing response tendencies in the state of uncertainty. The tree for new items allows the identification of new items as new ($DN$). Otherwise ($1-DN$), items were guessed to be old ($bO$) or new ($1-bO$). If guessed old, they were guessed as listened to ($bL$) or spoken ($1-bL$).

The model’s degrees of freedom (12) exceed the degrees of freedom in the data (10). To obtain a testable baseline model, three further parameters were restricted. $DOL$ was set equal to $DN$: item identification of listened-to items should be equal to the identification of new items (following Riefer et al., 2007). Guessing the correct person was restricted to $1/n = .20$. Again, a ceiling effect was obtained for reality monitoring that precluded interpretation of Hypothesis 3 (Experiment 1a: $dL = .99, SE = .01, CI [.97-.100], dS = .98, SE = .01, CI [.96-.99]$; Experiment 1b: $dL = .94, SE = .01, CI [.92-.97], dS = .95, SE = .01, CI [.93-.96]$). Therefore, we set equal the reality monitoring parameters for spoken and listened-to items ($dL = dS$). We used MultiTree for analyses (Moshagen, 2010).

**Multinomial Modeling Analyses of Experiment 1a**

Parameter estimates of the baseline model are shown in Figure 2. This model fit the data, $G^2(10) = 2.05, p = .15$. Further parameters are shown in Table A1. Testing item memory, the probability of identifying old items as old was set equal for spoken and listened-to items ($DOS = DOL$). This led to a significant loss of fit ($\Delta G^2[1] = 75.22, p < .001$). As expected, the recognition of spoken items exceeded the recognition of listened-to items, supporting Hypothesis 1. *Partner identification* was tested by setting the parameters for
partner identification equal for listened-to and spoken items ($pL = pS$). The significant loss in model fit ($\Delta G^2[1] = 97.10, p < .001$) indicates reduced partner identification when items were spoken rather than listened to, corroborating Hypothesis 2 of lower destination than source memory.

The response bias parameter ($bL$) indicates a tendency to guess items to be listened to ($bL$) when reality monitoring fails. Guessing would be by chance if $bL = 1-bL = .5$. Response bias significantly exceeded chance ($bL = .5: \Delta G^2[1] = 15.61, p < .001$), in line with the predicted response bias toward guessing “listened to”, corroborating Hypothesis 4. We refrained from a multinomial modeling analysis regarding conversational partner’s gender since several cells contained only 0-4 data points.

**Multinomial Modeling Analyses of Experiment 1b**

The baseline model fit the data, $G^2(10) = 0.05, p = .82$. Again, *item memory* was significantly better in spoken compared to listened-to items, corroborating Hypothesis 1 ($\Delta G^2[1] = 127.37, p < .001$). Corresponding to Experiment 1a, *partner identification* was significantly reduced in spoken compared to listened-to items ($\Delta G^2[1] = 135.70, p < .001$), again indicating better source than destination memory, in line with Hypothesis 2. As shown in Figure 2, replicating the ANOVA, partner identification in mixed-gender groups was better than in same-gender groups both for items spoken and listened to. As in Experiment 1a, Hypothesis 4 was supported by a significant *response bias* toward guessing listened to ($\Delta G^2[1] = 17.21, p < .001$).

**Discussion**

The aim of the present research was to test whether source memory exceeds destination memory in real-world interaction situations. At the same time, in line with the literature on generation, production, and enactment effects, spoken information should be remembered better than information listened to. Each of our participants spoke and listened to other
participants in a setup similar to speed dating. Across two experiments, both an ANOVA and a multinomial modeling analysis corroborated the hypothesis that spoken items were recognized better than listened-to items during a surprise recognition test. At the same time, findings converged on better partner identification for items listened to as opposed to spoken. The reality-monitoring hypothesis that identifying spoken items as spoken is superior to identifying listened-to items as listened to could not be confirmed, and reality monitoring in mixed-gender dyads was not superior to that in same-gender dyads. We replicated a typical guessing bias: Items not remembered were guessed to have been listened to. (The fewer) male partners were identified better than female partners, and overall partner identification in mixed-gender groups (Exp. 1a) was better than in same-gender groups (Exp. 1b). We first discuss each of these findings and then turn to their implications for conversation situations.

The present research provides an original contribution by creating a setting that allowed close experimental control over the study and test materials and high relevance to applied settings at the same time (also see Macrae et al., 1999; Perfect et al., 2009). We conceptually replicated the finding that source memory is better than destination memory (e.g., Gopie, Craik, & Hasher, 2010; Gopie & MacLeod, 2009; but see Lindner et al., 2014), thus attesting to its relevance for everyday interactions. Moreover, this finding is well in line with findings on negative effects of active involvement in external source monitoring (Jurica & Shimamura, 1999; Mulligan, 2004; 2011; Riefer et al., 2007). Our data further hint at a negative production effect (see MacLeod et al., 2010) in external source monitoring. But since we did not disentangle production from generation, future research should do so.

Our findings are consistent with the attention hypothesis (Gopie & MacLeod, 2009) and with the trade-off hypothesis (Jurica & Shimamura, 1999; Koriat et al., 1991; see Mulligan, 2011, for a more differentiated view on context memory). We have demonstrated that these hypotheses can be corroborated under ecologically more valid conversation
conditions. Our findings regarding better item memory for spoken than for listened-to items extend positive effects of active involvement (e.g., Jurica & Shimamura, 1999; Koriat et al., 1991; MacLeod et al., 2010) to dyadic interaction situations (also see Knutsen & Le Bigot, 2014). Thus, they corroborate the theoretical positions that active involvement in information output increases memory as compared to more passive encoding conditions (e.g., McDaniel & Bugg, 2008; Riefer et al., 2007; MacLeod et al., 2010; but see Steffens & Erdfelder, 1998, for boundary conditions).

We found no evidence for the reality monitoring hypothesis that identifying spoken items as spoken was superior to identifying listened-to items as listened to (for in-depth discussion, see Riefer et al., 2007). In addition to a ceiling effect, a possible reason for this is our omission of participants’ individual responses during the memory test that made the test a recognition of the generate cues, not the self-generated responses (e.g., Greenwald & Johnson, 1989). Whereas we did not counterbalance targets and distractors, given random assignment of many items to lists, it appears unlikely that this compromised our findings.

Experiment 1a used gender-mixed groups of conversation partners, whereas Experiment 1b relied on same-gender groups. Experiment 1a showed an effect of conversational partner’s gender on source and destination memory. Male partners, who were the minority, were identified better than female partners. This effect of partner gender was similar in size in the source and destination conditions. It thus appears that partner distinctiveness affected destination and source memory to similar degrees. Both this finding and the finding that destination memory, although significantly lower than source memory, was still good, demonstrate successful context integration. Whereas concentrating on what one is doing reduces context integration (Koriat et al., 1991), it does not eliminate it. Partner identification was generally better in Experiment 1a compared to Experiment 1b since the group of participants was more heterogeneous in the first case (e.g., Klauer & Wegener,
1998; Macrae et al., 1999): Even though gender is but one relevant social category in this respect, it is a salient one.

In the absence of memory for a given piece of information, we found a strong guessing bias that the item was listened to instead of spoken, replicating the “it had to be you” effect (e.g., Johnson & Raye, 1981; Ozubko et al., 2012). However, other experimental paradigms provoked what we call “it had to be me” effects (e.g., Lindner, Echterhoff, Davidson, & Brand, 2010; Macrae et al., 1999; Perfect et al., 2009). Thus, it is an open question which effect is more likely in applied settings. Possibly, the meta-memory heuristic that active involvement provokes good memory underlies these effects: If memory appears bad, “it had to be you”, if memory appears good, “it had to be me”. Future research delineating the conditions under which one or the other effect occurs could profit from using multinominal modeling.

Multinomial models are superior to ad hoc data correction methods for disentangling memory from guessing processes. The present model corroborated that destination memory is poorer than source memory when guessing biases are taken into account. We could not use an existing model that has been validated parameter by parameter because the present research integrates previous approaches to measuring source memory with those using the “Who-Said-What?” paradigm, where more than two sources are to be distinguished (e.g., Klauer & Wegener, 1998; Rakić, Steffens, & Mummendey, 2011). From the model used in the latter research, we adapted the technique to evaluate only whether phrases were assigned to the correct person or not, neglecting the different options of assigning a phrase to a wrong speaker/listener. We therefore believe this straightforward extension of the models proposed and validated by Meiser and Bröder (2002) as well as by Riefer and colleagues (2007) can be used without conducting an additional validation program of research.
Our results extend previous destination memory findings towards real world applications. Since conversations usually consist of speaking and listening to information, the within-subject design we used comes closer to natural situations. Additionally, destination memory was poorer than source memory also when sources/destinations were presented concurrently with items and when differential context cues as to the specific location in the experimental room could have been generated at test. Also, the finding held up when conversational features such as “minimal responses” and subtle nonverbal feedback were present that could have strengthened destination memory. However, this is not to say that our paradigm mimics natural conversations in all respects. For example, people do not only say what they have in mind, they also create elaborating thoughts; and they think about what to say next (e.g., Perfect et al., 2009). Actual phrase completions were never presented during recognition. Therefore, our findings should be extended to less restricted and controlled conversation situations. Also it is possible that recall of partner statements is higher than recall of own statements in get-acquainted conversations, where the focus is on learning about the partner (e.g., Stafford & Daly, 1984). And conversation topic could modify findings because information about oneself is overlearned, so participants may lose track what they have told their partners in the respective session or someone at another occasion (see Koriat & Ben-Zur, 1988; Brown, Hornstein, & Memon, 2006). In problem-solving conversations, where participants actively generate information on new topics, more re-use and higher recall of own statements compared to partners’ statements was found (e.g., Knutsen & Le Bigot, 2014; Miller, deWinstanley, & Carey, 1996). Additionally, the present paradigm should be extended to samples with weaker memory (strategies) including children, older people, or those with memory impairments (e.g., Gopie et al., 2010; El Haj, Postal, & Allain, 2013). Finally, it is yet unknown which interventions increase destination memory in
real conversations, for example, focusing on the partner (Gopie & MacLeod, 2009) or simply giving participants a learning instruction (Lindner et al., 2014).

As Lindner and colleagues (2014) elaborated, there are many situations in everyday life where source and destination memory are relevant (e.g., “Was it one of my dinner guests who told me they were vegetarian?”, “Have I told the cook about my food allergy?”). The present work has several implications for application. We have shown (once more) that active involvement in information output increases remembering the information. Thus, if one wants to remember a conversation topic, it is a good idea to actively participate in the conversation instead of listening only. Among mnemonic strategies, generating information and saying it out loud should therefore play a prominent role. Conversely, if there is disagreement whether something has been said or not, for example, during a work meeting, the probability is higher that the speaker correctly remembers it than any other person. Whereas speaking positively affects remembering content, we have also shown that it negatively affects remembering to whom something was said. Thus, speakers need to make an extra effort to remember to whom they said something; otherwise, the audience has an edge over them regarding this aspect. Whereas the speaker’s advantage in remembering the information itself is rather small (estimated around 10% in our experiments), the listener’s advantage in identifying the specific conversation partner may be quite substantial; for example, it was estimated to be around 30% better in the multinomial model in Experiment 1b. As this demonstrates, “Have I told you this already?” can be a substantial problem in communication even among young participants used to studying daily, and not only at conferences. As the comparison of our two experiments shows, remembering to whom one said something is most difficult if conversation partners are rather similar, for example, regarding their gender. Possibly, this similarity extends to other aspects of conversation partners. Thus, it would be more difficult to decide whether one has told something to one or
the other of one’s best friends, or to one or the other of one’s colleagues. Keeping track of
destinations of given information is essential in various situations, for example at school or at
work. A supervisor should know to whom of her employees she delegated which task. Failing
to remember the person in charge could unnecessarily consume resources or induce a
diffusion of responsibilities. Our findings underline the importance of taking notes and
writing protocols when several tasks are delegated. Possibly, destination memory problems
are augmented if some-one has addressed a whole group, being able to keep less track of who
was present or absent than in a case where attention was focused on an individual listener.

Retelling the same story twice may not only occur due to poor destination memory, but
also due to poor output monitoring, that is: Telling a story twice because of forgetting one
told it already (Brown et al., 2006). Additionally, it could be a case of unconscious
plagiarism (e.g. Macrae et al., 1999; Perfect et al., 2009): Unconsciously repeating oneself or
repeating information someone else said without being aware of this mistake. The integration
of output monitoring, destination memory, and unconscious plagiarism appears to be a
promising and comprehensive approach for investigating the question of why people repeat
themselves to others. According to target monitoring (Brown et al., 2006), people have to
keep track whether they have already given a specific piece of information to others or not
(output monitoring) and, additionally, to whom it has been provided (destination memory). In
addition to poor destination memory, the probability of retelling information to the same
person is regarded as high because people frequently interact with the same close others
(Brown et al., 2006). Additionally, actively thinking about others’ statements and ideas
increases the sense that the ideas actually belong to oneself, still increasing the chance of
retelling them in social contexts (Perfect et al., 2009). Thus, poor destination memory
appears to be a relevant but not sufficient explanation to resolve the question of why people
repeat themselves.
References


of Experimental Psychology: Learning, Memory, and Cognition, 28, 116–137. doi:
10.1037/0278-7393.28.1.116


Appendix

Table A1:

Further Parameter Estimates for Experiments 1a and 1b (with Standard Errors, SE, and Confidence Intervals, CI)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate (SE, CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Experiment 1a</td>
</tr>
<tr>
<td>$DN$</td>
<td>.89 (.01, .88 - .91)</td>
</tr>
<tr>
<td>$bO$</td>
<td>.13 (.02, .09 - .16)</td>
</tr>
<tr>
<td>$dL = dS$</td>
<td>.98 (.00, .97 - .99)</td>
</tr>
<tr>
<td>$eS$</td>
<td>.47 (.13, .20 - .73)</td>
</tr>
<tr>
<td>$eL$</td>
<td>.90 (.31, 0.29 - 1.50)</td>
</tr>
</tbody>
</table>

Note. The parameter $DN$ is the probability to identify new items as new; $bO$ is the guessing bias for guessing unidentified items to be old. The parameter $d$ represents the probability of reality monitoring: identifying spoken items as spoken ($dS$) or listened-to items as listened to ($dL$); $dS$ and $dL$ were set equal within each experiment. The parameter $e$ represents partner identification in the absence of reality monitoring; $eS$ for spoken and $eL$ for listened-to items.
Figure 1: Multinomial processing tree model of source and destination memory. The rectangles on the left represent the item presentation: listened to, spoken, new. The rectangles on the right represent the observed responses. The ovals in between represent latent cognitive states that lead to the respective response category. The probabilities of the respective cognitive states are represented by the following parameters: $DOL =$ the probability of recognizing an old item that was listened to as old; $DOS =$ the probability of recognizing an old spoken item as old; $DN =$ the probability of identifying a distractor as new; $bO =$ probability of guessing an item or distractor as old; $dL =$ correctly assigning a listened-to item as listened to; $dS =$ correctly assigning a spoken item as spoken; $bL =$ probability of guessing that a certain statement was listened to; $1-bL =$ probability of guessing that a certain statement was spoken; $pL =$ probability of identifying the correct partner when the item was listened to if reality monitoring was correct; $pS =$ probability of identifying the correct partner when the item was spoken if reality monitoring was correct; $eL =$ probability of identifying the correct partner when the item was listened to in the absence of reality monitoring; $eS =$ probability of identifying the correct partner when the item was spoken in the absence of reality monitoring; $1/n =$ probability of guessing the correct partner, where $n$ is the number of conversation partners.
Figure 2: Estimates of item memory, partner identification, reality monitoring, and guessing bias, separately for items that were spoken and items that were listened to in Experiments 1a and 1b. Reality monitoring parameters were set equal across speaking and listening, and guessing bias parameters linearly depend upon each other. Error bars show 95% confidence intervals.
Table 1:

*Sum Scores of Responses Aggregated over all Participants of Experiment 1a (upper part) and Experiment 1b (lower part).*

<table>
<thead>
<tr>
<th>Response Category</th>
<th>Listened to</th>
<th>Spoken</th>
<th>New</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Partner</td>
<td>Partner</td>
<td>Partner</td>
</tr>
<tr>
<td>True</td>
<td>Correct</td>
<td>Incorrect</td>
<td>Correct</td>
</tr>
<tr>
<td></td>
<td>1381</td>
<td>126</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>16</td>
<td>15</td>
<td>1276</td>
</tr>
<tr>
<td></td>
<td>31</td>
<td></td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>1894</td>
<td>340</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td>46</td>
<td>46</td>
<td>1731</td>
</tr>
<tr>
<td></td>
<td>47</td>
<td></td>
<td>32</td>
</tr>
</tbody>
</table>

*Experiment 1a*

*Experiment 1b*