Self-Repair Increases Referential Coordination

Gregory Mills,[a,b] Gisela Redeker[a]

[a] *Centre for Language and Cognition (CLCG), Faculty of Arts, University of Groningen*
[b] *School of Computer Science and Mathematics, Kingston University*

Abstract

When interlocutors repeatedly describe referents to each other, they rapidly converge on referring expressions which become increasingly systematized and abstract as the interaction progresses. Previous experimental research suggests that interactive repair mechanisms in dialogue underpin convergence. However, this research has so far only focused on the role of other-initiated repair and has not examined whether self-initiated repair might also play a role. To investigate this question, we report the results from a computer-mediated maze task experiment. In this task, participants communicate with each other via an experimental chat tool, which selectively transforms participants' private turn-revisions into public self-repairs that are made visible to the other participant. For example, if a participant, A, types "On the top square," and then before sending, A revises the turn to "On the top row," the server automatically detects the revision and transforms the private turn-revisions into a public self-repair, for example, "On the top square umm I meant row." Participants who received these transformed turns used more abstract and systematized referring expressions, but performed worse at the task. We argue that this is due to the artificial self-repairs causing participants to put more effort into diagnosing and resolving the referential coordination problems they face in the task, yielding better grounded spatial semantics and consequently increased use of abstract referring expressions.

Keywords: Dialog; Miscommunication; Repair; Conventionalization; Collaborative reference

Correspondence should be sent to Gregory Mills, School of Computer Science and Mathematics, Kingston University, Penrhyn Road, Kingston Upon Thames, Surrey KT1 2, UK. E-mail: g.mills@kingston.ac.uk

1. Introduction

A central finding in dialogue research is that when interlocutors repeatedly describe referents to each other, they rapidly converge on a shared set of referring expressions (Clark & Wilkes-Gibbs, 1986; Krauss & Weinheimer, 1966), which become progressively systematized and abstract (Healey, 1997; see Table 1). This occurs for a wide range of referents, for example, when describing spatial locations (Castillo, Smith, & Branigan, 2019; Garrod & Doherty, 1994; Mills, 2014; Nölle, Fusaroli, Mills, & Tylén, 2020), geometric figures (Bangerter, Mayor, & Knutsen, 2020), bodily features (Tylén, Fusaroli, Smith, & Arnoldi, 2020), music (Healey, Swoboda, Umata, & King, 2007), conceptual structures (Schwartz, 1995), confidence (Fusaroli et al., 2012), temporal sequences (Mills, 2011; Verhoef, Walker, & Marghetis, 2016), and also when describing how to manipulate physical objects (Shirozou, Miyake & Masukawa, 2002).

Convergence on referring expressions is intrinsically interactive. Each pair of participants typically creates their own, idiosyncratic conventions for the same referents depending on their specific interaction history (Garrod & Doherty, 1994; Healey, 1997). Yet, the development of abstraction is not simply due to the coordination problem of creating a novel referring expression: once referring expressions have been used successfully, they continue to develop in predictable directions (Garrod, 1999; Healey, 2004; see also Table 1). The emergence of abstract referring expressions also occurs across modalities: in spoken interaction (Pickering & Garrod, 2004), text-based messaging (Healey & Mills, 2006), gesture (Macuch Silva, Holler, Ozyurek, & Roberts, 2020; Motamedi, Schouwstra, Smith, Culbertson, & Kirby, 2019; Nölle, Staib, Fusaroli, & Tylén, 2018), whistle-based language (Verhoef, Roberts, & Dingemanse, 2015), and in graphical, mediated interaction (Galantucci, 2005; Healey,

Table 1
Global development of abstract descriptions in the maze game

Initially	You need to go to my switch which is all the way at the top right on the sticking out part on the left
5 min	That's me done, can you go two down from the large block of squares on the right
10 min	You need to go to the middle column last square
15 min	I'm on the fourth column from right 3rd square
20 min	Wait then go to 5th column topmost square
30 min	Went back to 4th column 1st square
35 min	3rd col from left, row 7 from top
40 min	Then 2c r 6. yours?
45 min	5, 7

Note. Initially, participants use descriptions that typically rely on visually salient features of the maze, for example, the "sticking out part" or "large block of squares on the right." As the task progresses, participants develop more systematized descriptions which conceive of the mazes as consisting of squares aligned in columns (e.g., "fourth column from right 3rd square"). By the end of the experiment, the most coordinated pairs tend to use extremely concise Cartesian coordinate descriptions which conceive of the mazes as consisting of rows and columns.

Swoboda, Umata, & Katagiri, 2001; Roberts, Lewandowski, & Galantucci, 2015; Theisen, Oberlander, & Kirby, 2010).

Further, the quality of the interaction directly affects the development of coordination. If interlocutors are prevented from providing each other with feedback, for example, by being prevented from annotating each other's drawings, this impedes the development of abstract referring expressions (Healey et al., 2007). Similarly, in multiparty interaction, convergence occurs at a different rate between fully ratified participants than between participants and overhearers who have limited opportunities for engaging in the interaction (Healey & Mills, 2006; see also Kühlen & Brennan 2013).

Cumulatively, these findings suggest that processing that occurs in interaction places important constraints on the semantic negotiation of referring expressions (Freyd, 1983; Olsen & Tylén, 2023). However, there is currently no consensus about which mechanisms are involved.

One important source of constraints comes from individuals' cognitive biases (Kirby, Cornish, & Smith, 2008). On this view, simply being exposed to another's linguistic output should suffice to drive abstraction, for example, in an iterated learning chain (Kirby, Griffiths, & Smith, 2014; Silvey, Kirby, & Smith, 2019). But when noninteracting participants are exposed to exactly the same signs as interacting dyads, the signs that are subsequently produced by the noninteracting participants are less effective and less efficient (Fay, Walker, Swoboda, & Garrod, 2018), demonstrating the importance of interindividual, as opposed to intraindividual processes occurring in interaction.

A parsimonious account for interindividual coordination is provided by the Interactive Alignment model (Pickering & Garrod, 2004), which proposes that convergence arises as a consequence of automatic mutual priming. But this does not fully explain convergence, since priming is intrinsically conservative (Healey, 2004): once a particular form is the most successfully and widely used by a group, there is no mechanism to explain how it might be supplanted by another. Yet, interlocutors do *not* settle on abbreviated forms of the initially most frequently used referring expression in a "winner-takes-all" process. Interlocutors continue to develop novel and more abstract descriptions throughout the interaction (see Table 1). The priming account also does not explain conversational routines that do not involve lexical repetition or syntactic parallelism, for example, adjacency pairs (Clark, 1996; Schegloff, 2007), which often consist of complementary pairs of *different* types of contribution. In fact, patterns of local imitation of turns are worse statistical predictors of dialogue coordination than patterns of different, complementary turns (Fusaroli & Tylen, 2016), while indiscriminate, local imitation is actually associated with unsuccessful dialogue (Fusaroli et al., 2012).

1.1. Miscommunication drives abstraction

An alternative account is provided by Healey (2008) and Healey, Mills, et al. (2018), who argue that the interactive mechanisms associated with miscommunication play a central role in the development of abstract descriptions. Although historically miscommunication has been treated as a phenomenon to be avoided by interlocutors (Healey, De Ruiter, & Mills, 2018), research in conversation analysis has revealed how miscommunication involves a family of intricate interactive "repair" mechanisms that are used by interlocutors to sustain

coordination (Dingemanse et al., 2015; Heesen, Fröhlich, Sievers, Woensdregt, & Dingemanse, 2022; Schegloff, 1992). For example,

Example 1

A	Hey, the first time they stopped me from selling cigarettes was this morning
B	From selling cigarettes
A	Or buying cigarettes

Source: From Schegloff et al. (1977).

Here, B supposes that A did not intend to say "selling," and identifies the possible mistake in A's turn by repeating the problematic phrase "From selling cigarettes." Importantly, this form of repair leaves it up to A to remedy their own mistake. Similar forms of repair occur in the maze game:

Example 2

A	Move to the third square second row
B	third?
A	from the right

In this example, participant B has trouble understanding how A is counting squares within a row. B signals this trouble by repeating the problematic element "*third*," which A then clarifies. Similarly, in Example 3:

Example 3

A	I'm in row 6 column 7
B	huh?
A	next to the clump of squares that looks like an arm sticking out

In this example, participant B is unable to precisely specify the problem. So, B uses an open-class repair, "*huh?*" (Drew, 1997), to signal problematic understanding, leading A to fully reformulate their turn with an easier to understand, less abstract, description.

According to the repair-based account of Healey et al., such repair sequences allow interlocutors to identify potential divergences of interpretation with their conversational partner concerning the semantics of referring expressions, and then interactively resolve these divergences. Findings from a set of maze-task experiments (Healey 1997; Healey & Mills, 2006; Healey et al., 2018; Mills, 2014; Mills & Healey, 2006) provide evidence for repair-driven convergence. In this task, pairs of participants collaboratively solve mazes. This presents participants with the recurrent need to refer to spatial locations (see Fig. 1 for an example of maze configuration). A consistent finding is that participants initially start out using descriptions which identify visually salient features of the maze, for example, "*the sticking out part*," "*at*
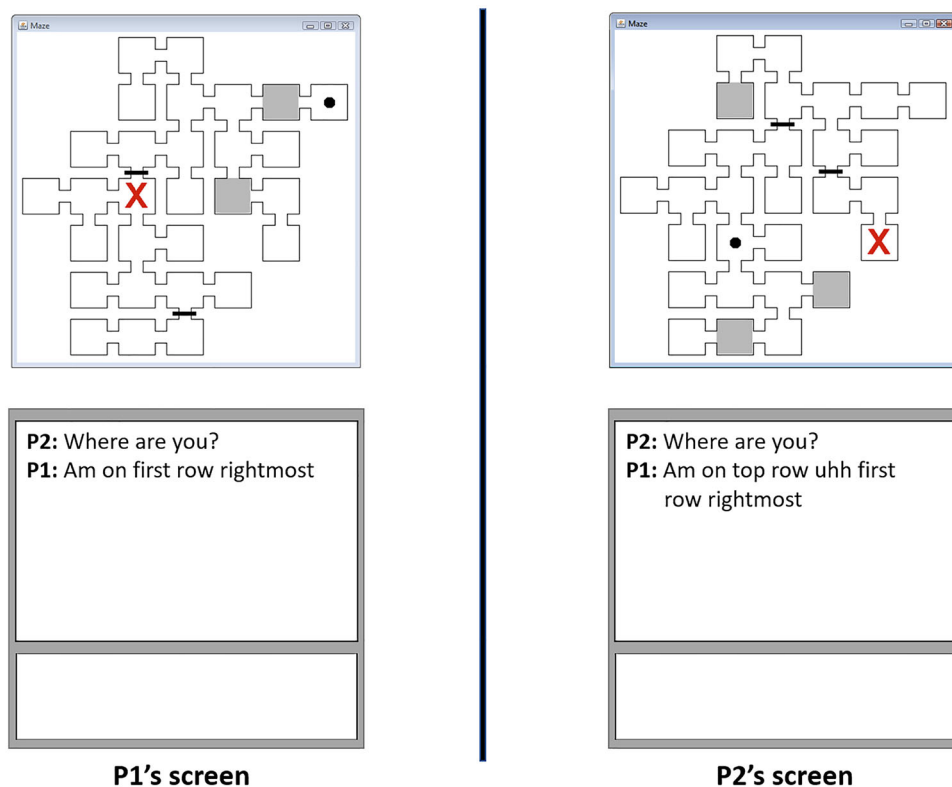
Fig. 1. Each participant has two windows on their screen. The top window displays the maze configuration. The lower window contains the chat interface used by participants to communicate with each other. In this dialogue, P1 originally typed "Am on top row," then subsequently deleted "top row" and replaced it with "first row." These private deletions are transformed by the server into a self-repair and sent to P2's screen.

the end of the arm." Over the course of the experiment, participants progressively use more abstract descriptions, for example, "*longest row 5th square*," while the most co-ordinated pairs converge on more complex abstract Matrix descriptions, such as "*A5*," "*2,1*," or "*row 3 column 4*" (see also Table 1). These descriptions are more difficult to coordinate on as their successful use requires coordinating on counting conventions (Healey, 2004). In a particular use, a description such as "*D4*" is a compact expression of a meaning like: "*4 across from the leftmost edge of the maze window, counting the edge as zero, and counting the missing boxes and counting 3 boxes up from the lower edge of the window*" (Mills, 2014). In order for participants to converge on such Matrix schemas, they first need to establish how to count rows and columns, whether to count from 0 or 1, whether "rows" can also be vertical, whether to count missing nodes, and so on (Healey, 2004). This is accomplished interactively, via repair, as it allows participants to identify, diagnose, and resolve any differences in interpretation (Healey, Mills et al., 2018; see also Woensdregt & Dingemanse, 2020; van Arkel, Woensdregt, Dingemanse, & Blokpoel, 2020; Bjørndahl, Fusaroli, Østergaard, & Tylén, 2015; Micklos, Walker, & Fay, 2020).

1.1.1. *Manipulating miscommunication*

To investigate experimentally the role played by repair, Mills and Healey (2006, 2008) conducted an experiment in which participants communicated via an experimental chat-tool which inserts artificial repairs into the interaction. The repairs appear, to participants, to originate from each other. In Examples 3 and 4 below, the second turn is an artificial repair produced by the server that appears to A as originating from participant B.

Example 4

A	Go to 3rd row 2nd column
B	row? *(produced by the server)*
A	yeah counting from the top

Example 5

A	My switch is at 4,5
B	huh? *(produced by the server)*
A	it's next to the sticking out part

Participants who received such artificial repairs produced fewer abstract descriptions, suggesting that, when participants encounter difficulties, they resort to less abstract descriptions that rely on visually salient features of the maze, which are easier to co-ordinate on.

A similar method was used in a subsequent experiment (Healey, Mills et al., 2018) which used the chat-tool to automatically detect instances of naturally occurring repair and amplify their severity. For example, in the following conversation, B's repair "*5th?*" is intercepted and transformed into "*what?*" and sent to A.

Example 6

A	go to the 3rd row 2nd column
B	3rd? *(intercepted by server, not sent to B)*
B	what? *(transformed turn sent to B)*
A	go to 3rd row 2nd column from the right

Participants who received these manipulations produced *more* abstract descriptions, while the manipulations had no other discernible effect on task performance. Healey, Mills et al. (2018) explain this pattern as being due to these interventions exacerbating the apparent severity of actual "trouble" in co-ordinating on the semantics of referring expressions. Participants respond to this increased severity by putting more effort into diagnosing and resolving the problem, yielding better grounded spatial semantics and consequently increased use of abstract descriptions.

1.2. Self-repair

In addition to the types of repair discussed above, speakers can also modify their own utterances with a so-called self-repair (Schegloff, 2007). Suppose that in Example 1, if A had corrected their slip of the tongue immediately after uttering it, this could have yielded a turn such as:

Example 7

A	Hey, the first time they stopped me from selling, uhh buying cigarettes was this morning

In this example, A identifies that they made a mistake, then signals the suspension of the delivery of the utterance with the editing expression "uhh" (Levelt, 1983), followed by replacing "buying" with "selling." Similarly, in the maze game, a participant might produce a self-repair such as:

Example 8

A	Move to the top row uhh I mean first row.

Example 9

A	My goal is at 4, 5 oops it's at 4, 6

From a cognitive perspective, self-repair can be attributed to the inexorable incrementality of processing (Gregoromichelaki, Kempson, & Howes, 2020; Hough, 2014). In addition, self-repairs are associated with better planning and coordination in effective team communication (Gervits, et al., 2016), are indicative of speakers adapting their descriptions to the perspective of their partner (Clark & Wilkes-Gibbs, 1986; Clark & Krych, 2004), and can have a beneficial effect on comprehension (Brennan & Schober, 2001).

1.3. Research questions

In summary, experimental research suggests that convergence on abstract referring expressions is underpinned by participants identifying, diagnosing, and resolving differences in interpretation via repair. However, experiments that manipulated repair have focused solely on *other*-initiated repair, thus perhaps missing how self-repair might be an important mechanism underpinning semantic change and adaptation in interaction. To address this question, we describe an experiment which investigates whether participants who play the maze task and whose (covert) self-repair efforts are artificially upgraded to public signals will also be induced to use more abstract descriptions.

2. Method

2.1. The maze task

The maze task is a computer-mediated version of the maze game experiments conducted by Garrod and Anderson (1987) and Garrod and Doherty (1994). Pairs of participants sit in different rooms in front of a computer screen which displays (1) the maze application and (2) a chat window for communicating with each other (see Fig. 1). The maze application displays a maze configuration consisting of interconnected nodes. Each participant's maze has a goal location marked with a red cross. The paths to the goal are blocked by gates which can only be opened if the other participant moves their position marker to a location that corresponds to a gray switch location that is only visible on their partner's screen. In order to get to the goal and solve the maze, participants have to open their gates by getting their partner to go onto a switch that only the player can see on their screen. This creates a recurrent co-ordination problem of participants guiding each other onto each other's switches. Participants play 12 randomly generated mazes, with a timeout of 5 min: If they fail to complete a maze in 5 min, the next maze is automatically loaded. This means that all dyads play 12 games (i.e., attempt 12 different mazes).

2.2. Manipulation: Transforming private turn-revision into self-repairs

Participants communicate with each other via a custom Artificial Intelligence-mediated (Hancock, Naaman, & Levy, 2020) instant messaging program (see Fig. 1). The instant messaging program consists of two windows. The top window shows the conversation history; the lower window is a turn-formulation window in which participants type their turn privately before sending it by pressing ENTER. All participants' keystrokes are sent to the server which analyses what they type and automatically transforms participants' private turn-revisions into self-repairs that are made visible to the other participant. For example, suppose a participant types the following:

Participant1: Go to the square on the left, next to the big block on top.

And then, before pressing ENTER, the participant edits the turn to:

Participant1: Go to the square on the left, next to the third column

The chat server automatically identifies the deleted portion of the turn, and appends an editing expression (Levelt, 1983), such as "umm," "uhh", followed by the new revised text. This would yield the following turn, sent to P2:

Participant1: Go to the square on the left, next to the big block on top umm next to the third column.

Importantly, participant 1 does not see the transformed text in their own chat window (see Fig. 1).

The experiment was conducted on native Dutch-speaking participants. We used the following editing expressions, which were identified in a previous pilot study: "*eeh*," "*eehm*,"

"*euh*," "*euhm*," "*ehm*," "*uh*," "*uuh*," "*uuhm*," "*ik bedoel*" ("*I mean*"), "*eh ik bedoel*" ("*uh I mean*"). Interventions were performed on both members of a dyad. The editing expression was selected randomly. In order to avoid cascades of interventions, a minimum of five turns had to elapse after each intervention before a turn by the same participant would be manipulated again.

2.3. Measures

The following measures were used:

Description types

Proportion of Matrix descriptions: This measure records whether a participant describes a Maze using a Cartesian coordinate schema consisting of rows and columns. Each turn was classified as one of three categories:

1. **Non spatial descriptions**, for example, "*tell me where to go*"
2. **Matrix**, Cartesian descriptions, for example, "*4,5*," "*A1*," "*row 3 column 2*"
3. **Other**, for example, "*the sticking out row*," "*the part that looks like a head*," "*big column on the right*." This corresponds to the categories Figural, Path, and Line from the original maze game (Garrod & Doherty, 1994).

Each maze description was classified independently by both authors. Any conflicting classification was discussed and resolved.

Performance measures

Task success: The number of mazes completed, which ranges between 0 and 12.

Number of turns: The number of messages typed in the private turn formulation window and sent to the other participant.

Turn-length: The length (in characters) of each message. Note that *turn-length* and *number of turns* measure different properties of the interaction. For example, if participants ground in multiple turns (installments), this would lead to more turns that are also shorter.

Edits: All turns were analyzed to establish whether they had been revised while being typed. This measures how much effort participants put into turn formulation.

Alignment: This records for each spatial description whether it is of the same type (Matrix vs. Other) as the description produced by the previous participant.

Additional analyses (see Discussion)

First use of Matrix description: This measure records when (i.e., on which turn number) a dyad first uses a Matrix description.

Unique words: This records the number of unique words produced by each participant.

2.4. Participants

One hundred and twenty-two participants were recruited from first-year undergraduate classes at the University of Groningen, Department of Communication and Information Science, and participated for course credit. These classes have an approximately equal balance of genders. Participants were randomly assigned to dyads, and the dyads were randomly assigned to either the *Control condition* or *Manipulated condition*. Four pairs were discarded as it turned out they had previously participated in a maze game experiment, leaving 24 dyads in the Control condition and 33 dyads in the Manipulated condition.

2.5. Procedure

Pairs were booked for 1-h slots. They were given written instructions, and then instructed verbally. Each pair of participants was asked to complete all 12 games as fast as possible. The nature of the experimental manipulations was not disclosed to participants until the debriefing session. All procedures were in accordance with the 1964 Helsinki Declaration and were reviewed by the Faculty's Committee for the Ethical Evaluation of Research (CETO).

3. Hypothesis and research question

3.1. Hypothesis

If repair underpins the emergence of abstract descriptions, then analogously to the experiment conducted by Healey, Mills, et al. (2018), participants whose covert repairs are exposed should use more Matrix descriptions than participants in the Control group.

3.2. Research question

What effect will the manipulation have on performance measures? We see three possibilities. The manipulation:

1. has no discernible effect (as in Healey, Mills, et al. 2018), or
2. increases the amount of "*trouble*" in the interaction, having a deleterious effect on task performance, or

3. increases participants' effort in coordinating in the task, having a beneficial effect on task performance.

4. Results

We analyzed the results using R version 3.6.2 (R Core Team, 2022), together with the LME4 package version 1.1-26 (Bates, Maechler, et al., 2015) and the MASS package v. 7.3-54 (Ripley et al., 2013). The models included random intercepts for dyads, participants, and mazes, as well as random slopes of condition and time within mazes. The models were estimated with an unstructured covariance matrix. Since we are interested in what the participants type, the artificial, transformed turns generated by the server are excluded from the analysis; only the original unmodified turns as sent by the participants are included in the analyses. This resulted in 17,627 turns overall. Dyads took a mean 31 min and 30 s (SD = 7 min and 4 s) to solve all 12 mazes.

4.1. Description types

In order to test whether participants in the Manipulated condition used more Matrix descriptions than participants in the Control condition, we conducted a likelihood ratio test of the model with the manipulation effect against the model without the manipulation effect. This revealed a significant difference ($\chi 2$ (3) = 8.52, p = .0364). The predicted probability of Matrix descriptions in the Control group is 0.02 [95% CI: 0.00, 0.18]. The predicted probability of Matrix descriptions in the Manipulated group is 0.34 [95% CI: 0.07, 0.77]. This confirms H1 (see Fig. 2).

4.2. Performance measures

In order to investigate the effect of the manipulations on performance measures, we compared models that included/excluded the corresponding predictors (main effects and interactions). Following Eshghi and Healey (2016) and Healey, Mills, et al. (2018), we pool the scores for the first six games (EARLY) and the last six games (LATE) to provide an index of how the measures change over time. Akaike's Information Criterion (AIC) was used for model comparison, as there was no nesting relationship between all models being compared—it was not possible to use a chi-square difference test between the different models. The model with the lowest AIC score was considered the best-fitting. We report the best-fitting model.

4.2.1. Task success

Task success was modeled with a multilevel binomial logistic regression, using glmer with a logit link function. The model with the lowest AIC showed a significant effect of the manipulation ($b = -0.939$ [95% CI: -1.64, -0.238], $z = -2.63$, $p < .01$) and an effect of time ($b = 3.66$ [95% CI: 1.85, 5.48], $z = 3.963$, $p < .001$) (see Fig. 3).
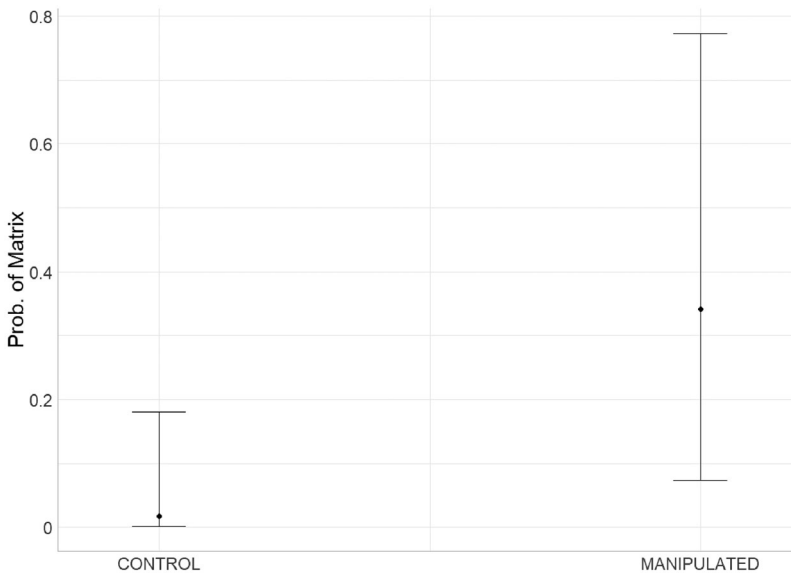
Fig. 2. Predicted probabilities of turns containing Matrix descriptions in Control and Manipulated dyads.
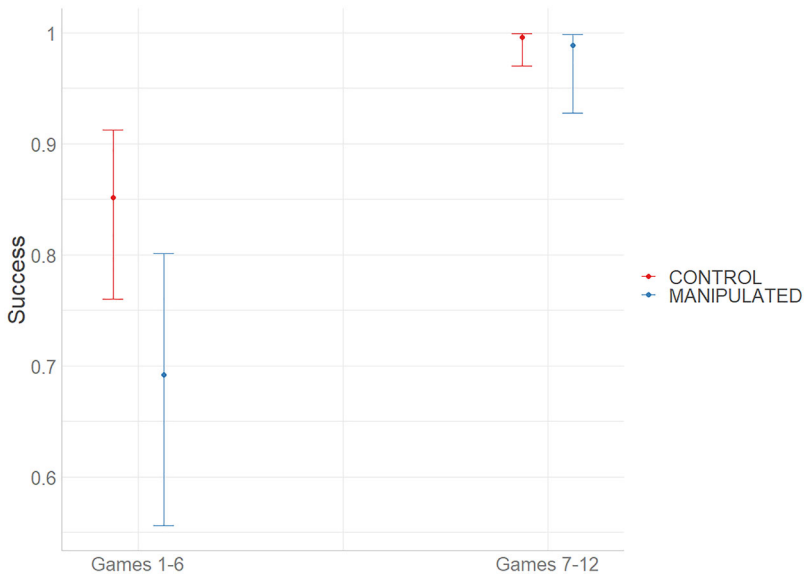

Fig. 3. Predicted probability of successfully solving a maze in Games 1–6 and 7–12.

4.2.2. Turn length

The length of participants' turns was modeled with a multilevel negative binomial regression, using glmer.nb with a log link function. The model with the lowest AIC showed a significant effect of time ($b = -0.210$ [95% CI: -2.61, -0.158], $z = -7.98$, $p < .001$).
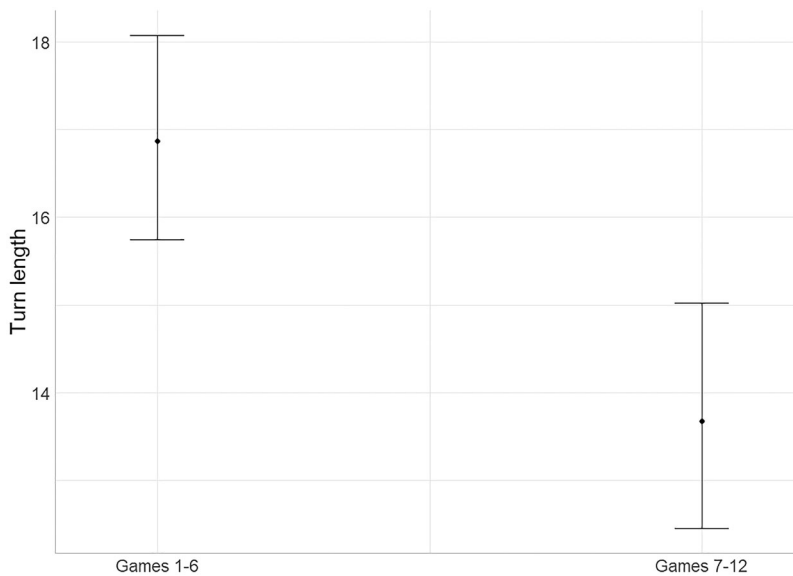
Fig. 4. Predicted turn length (characters) produced in Games 1–6 and last Games 7–12.

The predicted mean turn length in the first six games is 16.9 [95% CI: 15.8, 18.1] characters, while in the last six games, the predicted turn length is 13.7 [95% CI: 12.5, 15.0] characters (see Fig. 4).

4.2.3. Number of turns

The number of turns produced by dyads over the course of the experiment was modeled with a multilevel negative binomial regression, using glmer.nb with a log link function. The model with the lowest AIC showed a significant effect of the manipulation ($b = 0.187$, [95% CI: 0.00227, 0.372], $z = 1.984$, $p = .0473$) and a significant effect of time ($b = -0.512$, [95% CI: -0.593, -0.431], $p < .001$) (see Fig. 5 and Table 2).

4.2.4. Deletes

The number of deletes produced by dyads over the course of the experiment was modeled with a multilevel logistic regression, using glmer with a logit link function. The model with the lowest AIC showed a significant effect of time ($b = -0.211$, [95% CI: -0.308, -0.116], $z = -4.32$, $p < .01$). The predicted probability of a turn containing a delete is 0.39 [95% CI: 0.36, 0.42] in the first six games and is 0.34 [95% CI: 0.32, 0.37] in the last six games (see Fig. 6).

4.2.5. Semantic alignment

The alignment of participants' spatial descriptions was modeled with a multilevel logistic regression, using glmer with a logit link function. The model with the lowest AIC showed no significant effect of the manipulation ($b = -1.18$ [95% CI: -2.64, 0.284], $z = -1.68$,
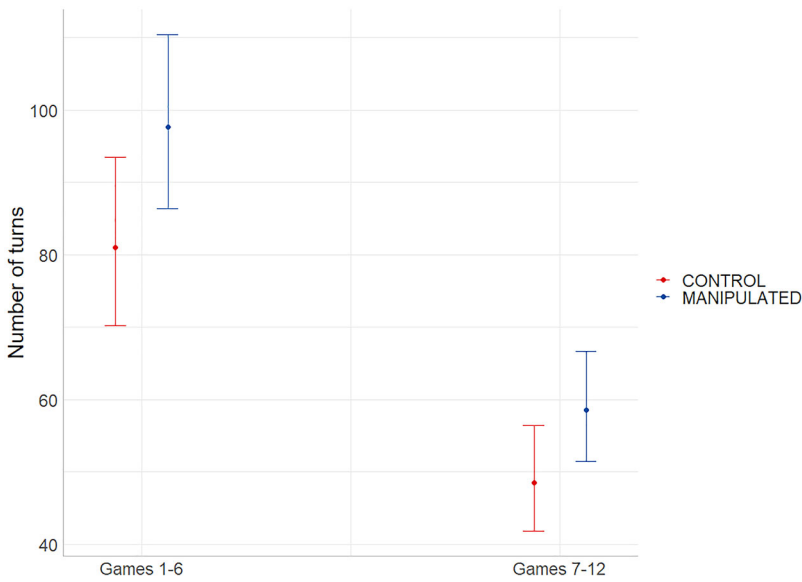
G. Mills, G. Redeker / Cognitive Science 47 (2023)

Fig. 5. Predicted number of turns produced in Games 1–6 and Games 7–12.

Table 2
Predicted number of turns

Games	Condition	Predicted number of turns	95% CI
1–6	Control	81	70.19, 93.47
7–12	Control	48.5	41.8, 56.4
1–6	Manipulated	97.7	86.4,110.5
7–12	Manipulated	58.5	51.4, 66.6

$p = .114$), but showed a significant effect of time ($b = 6.16$ [95% CI: 2.14, 10.2], $z = 3.002$, $p < .001$) and no significant interaction ($b = -1.23$, [95% CI: 4.19, 1.735], $z = 0.812$, $p = .417$). The predicted alignment score increases from 0.98 [95% CI: 0.94, 0.99] in the first six games to 1 in the last six games (see Fig. 7).

4.3. Additional measures

4.3.1. First use of matrix descriptions

The first use of Matrix descriptions by a dyad was modeled with a negative binomial regression from the MASS package. A likelihood ratio test of the model with the manipulation effect against the model without the manipulation effect did not reveal a significant difference ($\chi 2$ $(1) = 0.798$, $p = .372$). The predicted number of turns that elapse before a member of a dyad produces a Matrix description is 23.7 [95% CI: 16.5, 34.3].
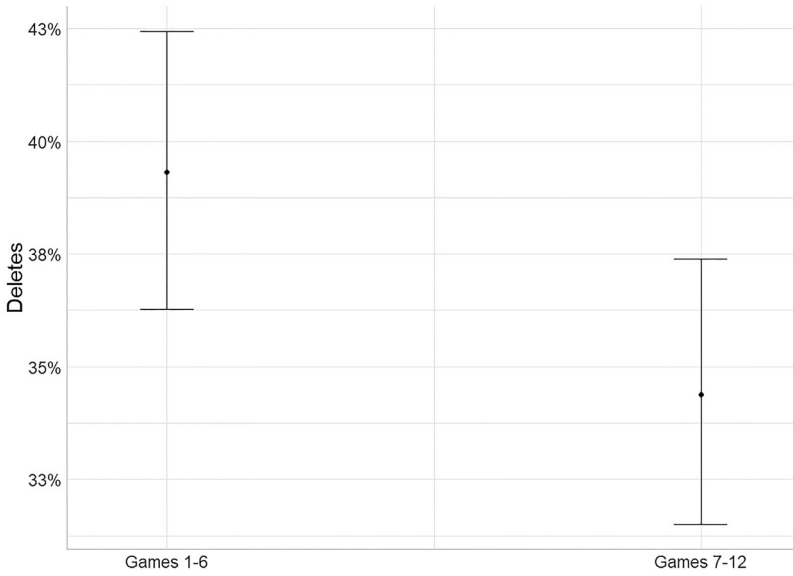
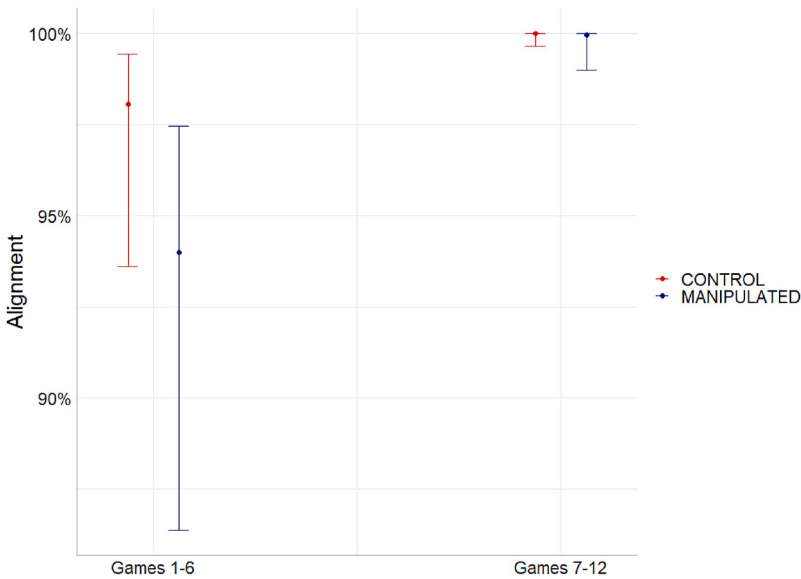Fig. 6. Predicted probability of a turn containing a delete.

Fig. 7. Predicted semantic alignment of participants' turns.

4.3.2. Number of unique words

The number of unique words was modeled with a multilevel negative binomial regression, using the glmer.nb with a log link function. The model with the lowest AIC showed a significant effect of the manipulation ($b = 0.129$ [95% CI: 0.0041, 0.255], $z = 2.02$,

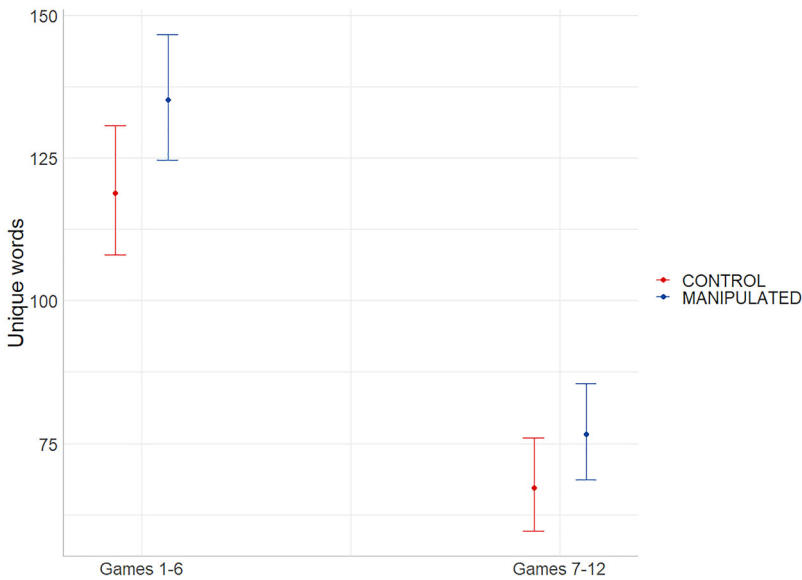*G. Mills, G. Redeker / Cognitive Science 47 (2023)*

Fig. 8. Predicted number of unique words in Games 1–6 and Games 7–12.

$p = .0429$) and also showed a significant effect of time ($b = -0.56$ [95% CI: -0.645, -0.492], $z = -14.6$, $p < .001$) (see Fig. 8 and Table 3).

5. Discussion

The results confirm the repair-driven view of co-ordination: Dyads whose covert repairs were exposed produced more abstract Matrix descriptions.

Although the changes in performance measures over time are consistent with the basic findings that interlocutors develop increasingly contracted referring expressions and become more successful as the task progresses, the effect of the manipulation on task performance is puzzling. Consistent with previous research, the manipulation appears to be having a beneficial effect on semantic co-ordination. However, unexpectedly, the manipulation also has a detrimental effect on task performance (task success and number of turns). Prima facie, this conflicts with previous research which has consistently found a positive association between abstract descriptions and task success (Castillo et al., 2019; Garrod & Doherty, 1994; Healey, Mills, et al., 2018).

The immediate questions that arise are: Why are the interventions causing more disruption? Why are they causing more abstraction? Might the increased abstraction be causing the disruption and/or vice-versa? We identify four possible explanations of the flow of information between participants:

(a) First, the *editing expressions* might be directly influencing participants to use more abstract descriptions. According to Arnold and Tanenhaus (2011), recipients of turns

Table 3
Predicted number of unique words

Games	Condition	Predicted unique words	95% CI
1–6	Control	118.8	108, 131
7–12	Control	67.3	59.6, 75.9
1–6	Manipulated	135.2	124.7, 146.6
7–12	Manipulated	76.6	68.6, 85.5

containing editing expressions are more likely to focus on less familiar items (see also Barr, 2001; Corley, MacGregor, & Donaldson, 2007). This could cause manipulated participants to consider previously unmentioned maze locations, effectively exposing participants to more exemplars, thereby creating a pressure to develop referring schemas that abstract over these exemplars (see also Raviv et al., 2022). Similarly, the editing expressions could also prompt participants to use less familiar referring expressions, stimulating participants' exploration of the space of possible descriptions. This is partially borne out in the additional analyses—manipulated dyads use more unique words, but do not appear to be introducing Matrix descriptions any earlier, suggesting that the decrease in task performance is not due to participants being induced to use Matrix descriptions prematurely, that is, before they have established sufficient co-ordination to use them successfully.

(b) Second, the *deleted text* might have influenced participants' conversational memory for the references produced during the interaction (Knutsen & Le Bigot, 2015) and could also have helped participants to uncover sources of misalignment (see also Schober, Suessbrick, & Conrad, 2018), in particular, concerning how to count in the maze. Consider Examples 10–15 below. In Examples 10 and 11, the deleted text shows that the sender is encountering trouble counting rows and nodes, potentially alerting the recipient to this trouble. Example 12 appears to show the sender was considering a different origo for counting (counting from bottom left vs. bottom right), while similarly Example 13 shows that the sender was originally considering counting from top to bottom, as opposed to from bottom to top. In Example 14, the edited turn refers to both horizontal and vertical row counts, whereas the original turn only shows horizontal row counts. Making the deleted text visible could be beneficial for the recipient since it provides an impetus for conceptualizing the maze as consisting of horizontal as well as vertical rows, which are the constituent elements of Matrix descriptions. Similarly, in Example 15, the sender initially conceptualized the maze as consisting of vertical rows, using ("*3rd row from left*") to refer to the vertical column which contains two switches and the position marker (see Fig. 1, P2's maze) but then changes and uses a different schema that conceptualizes the maze as consisting of horizontal rows ("*last row third block from left*").

Example 10:

Original text	drie naar beneden wandelen *go three down*
Manipulated text	dan vier of uh drie naar beneden wandelen *then go four or uh three down*

Example 11:

Original text	drie blokjes van links boven *three blocks from the left top*
Manipulated text	Twee blokjes van links boven ehm drie blokjes van links boven *two blocks from the left top ehm three blocks from the left top*

Example 12:

Original text	en twee blokjes van rechtsonder *and two blocks from the right bottom*
Manipulated text	en twee blokjes van link of euh rechtsonder *and two blocks from the left or uh right bottom*

Example 13:

Original text	Derde blokje van BENEDEn naar BOVEn *third block from bottom to top*
Manipulated text	Derde blokje va nboven nar uh van BENEDEn naar BOVEn *third block from top to uh from bottom to top*

Example 14:

Original text	3e rij horizontaal van onder *3rd row horizontal from bottom*
Manipulated text	3 verti eeh 3e rij horizontaal van onder *3rd vertical uh 3rd row horizontal from bottom*

Example 15:

Original text	Laatste rij derde blok van links *last row third block from left*
Manipulated text	3de rij van links of eehm Laatste rij derde blok van links *3rd row from left or uhm last row third block from left*

(c) Third, the editing expressions may have focused the recipients' attention on the immediately following words, that is, the revised text, thereby improving word-recognition (Brennan & Schober, 2001; Tree, 2001) and recall (Fraundorf & Watson, 2011; Corley et al., 2007).

(d) Fourth, the interventions could be causing participants to think their partner is experiencing more difficulty than they actually are, inducing them to compensate by expending more effort in grounding the referring expressions, following the "principle of least collaborative effort" (Clark & Brennan, 1991). This effect could be driven by the editing expressions, which have been shown to cause participants to appear less confident (Brennan & Williams,

1995) and as having a poorer grasp of the task (Susca & Healey, 2002). Moreover, many of the substitutions were concerned with correcting typos, for example

Example 16:

Original text	je hebt toch een verticale rij van 7 vakjes
	you do have a vertical row of 7 spaces
Manipulated text	je hebt toch een verticale rij van 7 bvalk of ehm vakjes
	you do have a vertical row of 7 bspac or uhm spaces

Example 17:

Original text	ik ben derde rij van boven helemaal links
	I am third row from top all the way on the left
Manipulated text	ik bern d uh ben derde rij van boven helemaal links
	I anm t uh am third row from top all the way on the left

In such manipulations, the text is much less readable than the original text and is often garbled. This could lead participants to think more negatively of their partner's ability (Boland & Queen, 2016), leading them to "dumb down" and put more effort into their turns in order to compensate for the (apparent) decreased skill (see, e.g., Dreisbach & Fischer, 2011) or commitment (Michael, Sebanz, & Knoblich, 2016; see also Mills, Gregoromichelaki et al., 2021) of their partner. In addition, participants' responses to manipulated messages could be giving the participant whose message was modified the impression that the recipient is experiencing more difficulty in formulating their turn, and is, therefore, a less credible interlocutor.

Relatedly, and more importantly in our view, the effect of the manipulation here is more radical than in previous experiments. Here, the manipulation renders as purposefully public information signals that were not intended to be included in the message sent to the interlocutor. Public self-repair in conversation does not only have the function of correcting trouble but can also be used strategically by the speaker to perform other actions like marking dispreferred responses, serve identity construction, or responsively react to (multimodal) feedback from the addressee (see, e.g., Lerner & Kitzinger, 2007; Schegloff, 2013). Under our manipulation, such potential strategic uses appear in the public arena for the consideration of the addressee while not underpinned by the intentions of the speaker or any reasons based on the interactional common ground, for example, there is no dispreferred response that is mitigated and the reformulations are not intended to indicate that the speaker's original description needs to be taken into account by the addressee. It is possible that such nonintended messages have both a local and downstream effect on the amount of effort that participants have to expend to disentangle what the import of each other's responses is. This might facilitate coordination in making the speaker's thought process transparent and liable to be corrected (see the discussion of editing expressions above), but, on the other hand, it might result in participants having to take more turns to achieve their goal.

5.1. Conclusions and future work

In summary, it appears that self-repairs are causing participants to put more effort into grounding their referring expressions, whether as a consequence of attributing lower confidence to the other interlocutor, or due to the edited text providing more information about problems in the task. Dyads who received the interventions typed more turns and solved fewer mazes, suggesting that they are putting more effort into co-ordinating their referring expressions, while the increased use of unique words suggests that the interventions are inducing participants to explore the space of possible referring expressions. Somewhat surprisingly, despite using more abstract descriptions, participants are not using them earlier, suggesting that the exploration process is not occurring at the level of Matrix descriptions, but is presumably occurring at a finer grain, for example, in clarifying spatial semantics, as in Examples 10–15.

However, it is unclear how the constituent components of the self-repairs contributed to the patterns observed. The interventions used a variety of editing expressions, which might have had different effects on participants (see, e.g., Clark & Fox Tree, 2002; Womack et al., 2012 for a discussion). Also, the algorithm for transforming private edits into public self-repairs was not sensitive to the content of the messages. This means that many different types of "trouble" were made visible, including typos, reformulations, and specifications. In addition, some of this "trouble" might have been introduced by the interventions themselves, as well as by participants attempting to make sense of their partner's response to the interventions. Given this complexity, it is very difficult to determine the extent to which the different types of editing expression and "trouble" types might have contributed toward the observed pattern.

To address these issues, a promising next step would be to use more sophisticated AI-mediated communication to detect and manipulate specific kinds of "trouble," for example, solely manipulating typos or reformulations of Matrix descriptions, as well as "trouble" concerning the procedural coordination in the task (Knutsen, Bangerter, & Mayor, 2019). In contrast to the present study, such experiments should use a between-participant design to avoid interactions between different types of intervention. This would allow much more fine-grained comparison of different types of self-repair, and would also allow comparison between self- and other-initiated repair: the same reformulations of referring expressions could be introduced in one dialogue in the format of artificial self-repairs, and in a different dialogue as artificial other-initiated repairs. This approach could be augmented by using an incremental WYSIWYG chat interface which displays characters as they are typed (Maraev, Mazzocconi, Mills, & Howes, 2020; Ziembowicz & Nowak, 2019), which would be amenable to artificially and automatically manipulating public turn-edits in real-time.

Data and analysis are stored here: https://osf.io/prdz5/

References

Arnold, J. E., & Tanenhaus, M. K. (2011). Disfluency effects in comprehension: How new information can become accessible. In Gibson, E. A., & Pearlmutter, N. J. (Eds.), *The processing and acquisition of reference* (pp. 197–217). MIT Press.

Arkel, J. V., Woensdregt, M. S., Dingemanse, M., & Blokpoel, M. (2020). A simple repair mechanism can alleviate computational demands of pragmatic reasoning: Simulations and complexity analysis. In *Proceedings of the 24th Conference on Computational Natural Language Learning*.

Bangerter, A., Mayor, E., & Knutsen, D. (2020). Lexical entrainment without conceptual pacts? Revisiting the matching task. *Journal of Memory and Language*, *114*, 104129.

Barr, D. J. (2001). Trouble in mind: Paralinguistic indices of effort and uncertainty in communication. In Cave, C., Guaitella, I., & Santi, S. (Eds.), *Oralité et gestualité: Interactions et comportements multimodaux dans la communication*, 597–600.

Bates, D., Machler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*, 1–48.

Bjørndahl, J. S., Fusaroli, R., Østergaard, S., & Tylén, K. (2015). Agreeing is not enough: The constructive role of miscommunication. *Interaction Studies*, *16*(3), 495–525.

Boland, J. E., & Queen, R. (2016). If you're house is still available, send me an email: Personality influences reactions to written errors in email messages. *PLoS one*, *11*(3), e0149885.

Brennan, S. E., & Schober, M. F. (2001). How listeners compensate for disfluencies in spontaneous speech. *Journal of Memory and Language*, *44*(2), 274–296.

Brennan, S. E., & Williams, M. (1995). The feeling of another's knowing: Prosody and filled pauses as cues to listeners about the metacognitive states of speakers. *Journal of Memory and Language*, *34*(3), 383–398.

Castillo, L., Smith, K., & Branigan, H. P. (2019). Interaction promotes the adaptation of referential conventions to the communicative context. *Cognitive Science*, *43*(8), e12780.

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

Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. In: *Perspectives on socially shared cognition*, 127–149. Washington, DC: American Psychological Association.

Clark, H. H., & Krych, M. A. (2004). Speaking while monitoring addressees for understanding. *Journal of memory and language*, *50*(1), 62–81.

Clark, H. H., & Tree, J. E. F. (2002). Using uh and um in spontaneous speaking. *Cognition*, *84*(1), 73–111.

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

Corley, M., MacGregor, L. J., & Donaldson, D. I. (2007). It's the way that you, er, say it: Hesitations in speech affect language comprehension. *Cognition*, *105*(3), 658–668.

Dingemanse, M., Roberts, S. G., Baranova, J., Blythe, J., Drew, P., Floyd, S., Gisladottir, R. S., Kendrick, K. H., Levinson, S. C., Manrique, E., Rossi, G., & Enfield, N. J. (2015). Universal principles in the repair of communication problems. *PLoS One*, *10*(9), e0136100.

Dreisbach, G., & Fischer, R. (2011). If it's hard to read… try harder! Processing fluency as signal for effort adjustments. *Psychological Research*, *75*, 376–383.

Drew, P. (1997). 'Open' class repair initiators in response to sequential sources of troubles in conversation. *Journal of Pragmatics*, *28*(1), 69–101.

Eshghi, A., & Healey, P. G. (2016). Collective contexts in conversation: Grounding by proxy. *Cognitive science*, *40*(2), 299–324.

Fay, N., Walker, B., Swoboda, N., & Garrod, S. (2018). How to create shared symbols. *Cognitive Science*, *42*, 241–269.

Fraundorf, S. H., & Watson, D. G. (2011). The disfluent discourse: Effects of filled pauses on recall. *Journal of memory and language*, *65*(2), 161–175.

Freyd, J. J. (1983). Shareability: The social psychology of epistemology. *Cognitive Science*, *7*(3), 191–210.

Fusaroli, R., & Tylén, K. (2016). Investigating conversational dynamics: Interactive alignment, interpersonal synergy, and collective task performance. *Cognitive Science*, *40*(1), 145–171.

Galantucci, B. (2005). An experimental study of the emergence of human communication systems. *Cognitive Science*, *29*(5), 737–767.

Garrod, S., & Anderson, A. (1987). Saying what you mean in dialogue: A study in conceptual and semantic co-ordination. *Cognition*, *27*(2), 181–218.

Garrod, S., & Doherty, G. (1994). Conversation, co-ordination and convention: An empirical investigation of how groups establish linguistic conventions. *Cognition*, *53*(3), 181–215.

Gervits, F., Eberhard, K., & Scheutz, M. (2016). Team communication as a collaborative process. *Frontiers in Robotics and AI*, *3*, 62.

Gregoromichelaki, E., Kempson, R., & Howes, C. (2020). Actionism in syntax and semantics. *Dialogue and Perception*, *CLASP Papers in Computational Linguistics, 2*, pp. 12-27

Hancock, J. T., Naaman, M., & Levy, K. (2020). AI-mediated communication: Definition, research agenda, and ethical considerations. *Journal of Computer-Mediated Communication*, *25*(1), 89–100.

Healey, P. G. (2004). Dialogue in the degenerate case? *Behavioral and Brain Sciences*, *27*(2), 201–201.

Healey, P. G. (1997). Expertise or expertese?: The emergence of task-oriented sub-languages. In *Proceedings of the 19th Annual Conference of the Cognitive Science Society*.

Healey, P. G., Swoboda, N., Umata, I., & Katagiri, Y. (2001). Representational form and communicative use. In *Proceedings of the 23rd Annual Conference of the Cognitive Science Society*.

Healey, P. G., & Mills, G. (2006). Participation, precedence and co-ordination in dialogue. In *Proceedings of the 28th Annual Conference of the Cognitive Science Society*. Vancouver: Cognitive Science Society.

Healey, P. G., Swoboda, N., Umata, I., & King, J. (2007). Graphical language games: Interactional constraints on representational form. *Cognitive Science*, *31*(2), 285–309.

Healey, P. G., De Ruiter, J. P., & Mills, G. J. (2018). Editors' introduction: Miscommunication. *Topics in Cognitive Science*, *10*(2), 264–278.

Healey, P. G., Mills, G. J., Eshghi, A., & Howes, C. (2018). Running repairs: Coordinating meaning in dialogue. *Topics in Cognitive Science*, *10*(2), 367–388.

Heesen, R., Fröhlich, M., Sievers, C., Woensdregt, M., & Dingemanse, M. (2022). Coordinating social action: A primer for the cross-species investigation of communicative repair. *Philosophical Transactions of the Royal Society B*, *377*(1859), 20210110.

Hough, J. (2014). Modelling incremental self-repair processing in dialogue (Doctoral dissertation), Queen Mary University of London.

Kirby, S., Cornish, H., & Smith, K. (2008). Cumulative cultural evolution in the laboratory: An experimental approach to the origins of structure in human language. *Proceedings of the National Academy of Sciences*, *105*(31), 10681–10686.

Kirby, S., Griffiths, T., & Smith, K. (2014). Iterated learning and the evolution of language. *Current Opinion in Neurobiology*, *28*, 108–114.

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*(2), 574.

Knutsen, D., Bangerter, A., & Mayor, E. (2019). Procedural coordination in the matching task. *Collabra: Psychology*, *5*(1), 3.

Krauss, R. M., & Weinheimer, S. (1966). Concurrent feedback, confirmation, and the encoding of referents in verbal communication. *Journal of Personality and Social Psychology*, *4*(3), 343.

Kuhlen, A. K., & Brennan, S. E. (2013). Language in dialogue: When confederates might be hazardous to your data. *Psychonomic Bulletin & Review*, *20*(1), 54–72.

Lerner, G. H., & Kitzinger, C. (2007). Extraction and aggregation in the repair of individual and collective self-reference. *Discourse Studies*, *9*(4), 526–557.

Levelt, W. J. (1983). Monitoring and self-repair in speech. *Cognition*, *14*(1), 41–104.

Maraev, V., Mazzocconi, C., Mills, G., & Howes, C. (2020). "LOL what?": Empirical study of laughter in chat based dialogues. In *Proceedings of the Laughter and Other Non-Verbal Vocalisations Workshop 2020, Bielefeld, Germany*.

Macuch Silva, V., Holler, J., Ozyurek, A., & Roberts, S. G. (2020). Multimodality and the origin of a novel communication system in face-to-face interaction. *Royal Society Open Science*, 7(1), 182056.

Michael, J., Sebanz, N., & Knoblich, G. (2016). Observing joint action: Coordination creates commitment. *Cognition*, *157*, 106–113.

Micklos, A., Walker, B., & Fay, N. (2020). Are people sensitive to problems in communication? *Cognitive Science*, *44*(2), e12816.

Mills, G. J. (2014). Dialogue in joint activity: Complementarity, convergence and conventionalization. *New Ideas in Psychology*, *32*, 158–173.

Mills, G. (2011). The emergence of procedural conventions in dialogue. In *Proceedings of the Annual Meeting of the Cognitive Science Society*.

Mills, G. J., & Healey, P. G. (2006). Clarifying spatial descriptions: Local and global effects on semantic coordination. In *Proceedings of SemDial*. Universität Potsdam.

Mills, G., & Healey, P. (2008). Semantic negotiation in dialogue: The mechanisms of alignment. In *Proceedings of the 9th SIGdial Workshop on Discourse and Dialogue*.

Mills, G., Gregoromichelaki, E., Howes, C., & Maraev, V. (2021). Influencing laughter with AI-mediated communication. *Interaction Studies 22*(3), 416–463.

Motamedi, Y., Schouwstra, M., Smith, K., Culbertson, J., & Kirby, S. (2019). Evolving artificial sign languages in the lab: From improvised gesture to systematic sign. *Cognition*, *192*, 103964.

Nölle, J., Fusaroli, R., Mills, G. J., & Tylén, K. (2020). Language as shaped by the environment: Linguistic construal in a collaborative spatial task. *Palgrave Communications*, *6*(1), 1–10.

Nölle, J., Staib, M., Fusaroli, R., & Tylén, K. (2018). The emergence of systematicity: How environmental and communicative factors shape a novel communication system. *Cognition*, *181*, 93–104.

Olsen, K., & Tylén, K. (2023). On the social nature of abstraction: Cognitive implications of interaction and diversity. *Philosophical Transactions of the Royal Society B*, *378*(1870), 20210361.

Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, *27*(2), 169–190.

R Core Team. (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Raviv, L., Lupyan, G., & Green, S. C. (2022). How variability shapes learning and generalization. *Trends in Cognitive Sciences*. *26*(6), pp. 462–483.

Ripley, B., Venables, B., Bates, D. M., Hornik, K., Gebhardt, A., Firth, D., & Ripley, M. B. (2013). Package 'mass'. *Cran r*, *538*, 113–120.

Roberts, G., Lewandowski, J., & Galantucci, B. (2015). How communication changes when we cannot mime the world: Experimental evidence for the effect of iconicity on combinatoriality. *Cognition*, *141*, 52–66.

Schegloff, E. A. (1992). Repair after next turn: The last structurally provided defense of intersubjectivity in conversation. *American Journal of Sociology*, *97*(5), 1295–1345.

Schegloff, E. A. (2007). *Sequence organization in interaction: A primer in conversation analysis I* (Vol. 1). Cambridge University Press.

Schegloff, E. A. (2013). Ten operations in self-initiated, same-turn repair. *Conversational repair and human understanding*, *30*, 41–70.

Schegloff, E. A., Jefferson, G., & Sacks, H. (1977). The preference for self-correction in the organization of repair in conversation. *Language*, *53*(2), 361–382.

Schober, M. F., Suessbrick, A. L., & Conrad, F. G. (2018). When do misunderstandings matter? Evidence from survey interviews about smoking. *Topics in Cognitive Science*, *10*(2), 452–484.

Schwartz, D. L. (1995). The emergence of abstract representations in dyad problem solving. *Journal of the Learning Sciences*, *4*(3), 321–354.

Shirouzu, H., Miyake, N., & Masukawa, H. (2002). Cognitively active externalization for situated reflection. *Cognitive Science*, *26*(4), 469–501.

Silvey, C., Kirby, S., & Smith, K. (2019). Communication increases category structure and alignment only when combined with cultural transmission. *Journal of Memory and Language*, *109*, 104051.

Susca, M., & Healey, E. C. (2002). Listener perceptions along a fluency–disfluency continuum: A phenomenological analysis. *Journal of Fluency Disorders*, *27*(2), 135–161.

Theisen, C. A., Oberlander, J., & Kirby, S. (2010). Systematicity and arbitrariness in novel communication systems. *Interaction Studies*, *11*(1), 14–32.

Tree, J. E. F. (2001). Listeners' uses of um and uh in speech comprehension. *Memory & Cognition*, *29*(2), 320–326.

Tylén, K., Fusaroli, R., Smith, P., & Arnoldi, J. (2020). The social route to abstraction: Interaction and diversity enhance rule-formation and transfer in a categorization task. Retrieved from https://psyarxiv.com/qs253

Verhoef, T., Roberts, S. G., & Dingemanse, M. (2015). Emergence of systematic iconicity: Transmission, interaction and analogy. In *37th Annual Conference of the Cognitive Science Society* (pp. 2481–2486). Cognitive Science Society.

Verhoef, T., Walker, E., & Marghetis, T. (2016). Cognitive biases and social coordination in the emergence of temporal language. In *Proceedings of the 38th Annual Conference of the Cognitive Science Society* (pp. 2615–2620). Cognitive Science Society.

Woensdregt, M., & Dingemanse, M. (2020). Modelling the role of other-initiated repair in facilitating the emergence of compositionality. In *The Evolution of Language: Proceedings of the 13th International Conference (EvoLang13)*.

Womack, K., McCoy, W., Alm, C. O., Calvelli, C., Pelz, J. B., Shi, P., & Haake, A. (2012, July). Disfluencies as extra-propositional indicators of cognitive processing. In Proceedings of the workshop on extra-propositional aspects of meaning in computational linguistics (pp. 1–9).

Ziembowicz, K., & Nowak, A. (2019). Prosody of text communication? How to induce synchronization 3 and coherence in chat conversations. *International Journal of Human–Computer Interaction*, *35*(17), 1586–1595.