

Cognitive and social delays in the initiation of conversational repair

Julia Mertens

Tufts University, Medford, MA, 02155

JULIA.MERTENS@TUFTS.EDU

Jan P. de Ruiter

Tufts University, Medford, MA, 02155

JP.DERUITER@TUFTS.EDU

Editor: Patrick G.T. Healey

Submitted 04/2020; Accepted 01/2021; Published online 03/2021

Abstract

The exact timing of a conversational turn conveys important information to listeners. Most turns are initiated within 250ms after the previous turn. However, interlocutors take longer to initiate certain types of turns: those that either require more cognitive processing or are socially dispreferred. Many dispreferred turns are also cognitively demanding, so it is difficult to attribute specific conversational delays to social or cognitive mechanisms. In this paper, we evaluate how cognitive and social variables contribute to the timing of utterances in conversation. We focus on a type of turn that is socially dispreferred, cognitively demanding, and generally delayed: other-initiations of repair (OIRs). OIRs occur when a listener notices and decides to signal a comprehension problem (e.g., “What?”). We analyzed the Floor Transfer Offsets of 456 OIRs. Interlocutors initiated OIRs later when trouble source turns had weaker discourse context or were shorter. We found cognitive effects of trouble source duration and discourse context: a longer duration or stronger context was associated with shorter OIR FTOs. We also found social effects of problem attribution and size: when the problem could be attributed to the environment or was smaller, OIR FTOs were shorter. Discourse context, planning, and social attribution manifest in the timing of turns.

Keywords: Other-initiated repair, prediction, floor transfer offset, preference organization, turn taking

1. Introduction

Humans develop and maintain relationships through conversation. In fact, conversation is ubiquitous: children acquire language by interacting with their parents and peers, and coworkers increase their social coherence by chatting around the water cooler. However, conversation is not as effortless as it seems. On the contrary, conversation is an advanced cognitive skill. Interlocutors must perform many tasks at once. Speakers coordinate their words, prosody, and other non- and para-verbal signals to form messages tailored to their recipient and the context. Listeners integrate that information into their discourse record as they plan their response. At the same time, all interlocutors update their common ground, which involves accessing both their short-term and long-term memory. To perform all these tasks at once, conversants draw heavily on cognitive resources.

These taxing cognitive tasks are also time sensitive. Interlocutors aim to minimize gaps and overlaps (Sacks et al., 1978). As a consequence, speakers usually begin their turn within 250ms of

the end of the previous turn (de Ruiter et al., 2006; Stivers et al., 2009). In comparison, participants in psycholinguistic experiments take more than twice as long (according to Indefrey and Levelt, 2004, approximately 600ms) to initiate the name of a simple object. It seems much more complex to plan a conversational turn than to plan the name of an object. Nevertheless, interlocutors initiate conversational turns much earlier than the processing time for simple object naming would suggest.

Turn transition times in conversation are so short because listeners predict the upcoming turn (de Ruiter et al., 2006; Magyari and de Ruiter, 2012; Magyari et al., 2017) and plan their response before the current turn ends (Jansen et al., 2014; Magyari et al., 2017; Bögels et al., 2020). Even so, speakers initiate some turns later than others. In fact, speakers consistently delay some types of turns (Schegloff, 1979). For example, people usually pause before declining (but not accepting) invitations.

Psycholinguists and conversation analysts explain conversational delays differently. Psycholinguists argue that speakers initiate turns later when they need to process or manipulate more information. Processing takes time; more processing takes more time (Donders, 1969). For example, interlocutors initiate lies later than truths (Walczyk et al., 2003) because lies require additional cognitive processes: deciding to lie, inhibiting the truth, and constructing the lie (Walczyk et al., 2003). From the psycholinguistic perspective, delays are not part of the speaker's intended message. Instead, delays are unintentional, automatic side effects of increased cognitive processing; any time the listener's processing load increases they take longer to initiate their response.

Conversation analysts argue that some delays reflect *preference organization* and speakers delay turns to signal that the turn will be *dispreferred* (Clayman, 2002). Dispreferred responses do not follow the current action of the conversation or do not create social solidarity (Schegloff, 1968). In contrast, speakers do not delay *preferred* turns, which continue the current action of the conversation or create social solidarity. In Excerpt 1, Les attempts to compensate Joy for a favor. On line 4, Joy rejects Les's offer. Rejections are typically dispreferred turns, but it is preferred to reject offers. The rejection overlaps with the offer. The timing of the rejection emphasizes that the offer is unnecessary. Early responses convey that the response should be expected or that the speaker is eager to produce the utterance.

Excerpt 1: Fast rejection on line 4, from Kendrick and Torreira (2015)

1 Les: .hhh P'APS you'd like- Would you like eh:m:: some
 2 frozen f::: # ruit fr'm our k- ou:r # freezer as
 3 a small recompens[e?
 4 **Joy:** [Oh: **Les for goodness sake n:o**
 5 **I don't want anything**

If a speaker does not delay a dispreferred turn, they risk appearing rude. Excerpt 2 is a fictional, modified version of Excerpt 1. Joy's response (line 2-3) now rejects an invitation. The rejection implies an eagerness to reject Les, or that Les should have expected a rejection. There may be some unknown negative history between Les and Joy, or Joy is rude. Speakers may delay initiating a rejection to avoid conveying an inappropriate sentiment.

Excerpt 2: Fictional rejection on line 3

1 Les: .hhh P'APS you'd like to go out to dinn[er?
 2 **Joy:** [Oh: **Les**
 3 **for goodness sake n:o I don't**

To summarize, delays in initiating conversational turns could be due to cognitive variables, social variables, or a combination of the two. In this study, we examine social and cognitive variables together to determine their relative contributions to delays in conversation. Roberts et al. (2015) analyzed how both cognitive processing and preference organization influenced the timing of all utterances in a corpus. However, different types of utterances have different timing mechanisms. For example, backchannels (e.g., “uh-huh” or “yeah”; Schegloff, 1982; Yngve, 1970), insults, and questions occur in different contexts, perform different functions, and involve different social and cognitive information. To avoid these confounds, we focus on one type of utterance that (1) requires greater levels of cognitive processing, (2) is socially dispreferred and (3) tends to be delayed in comparison to other turns. Together, these features suggest that the utterance is delayed because it requires increased cognitive processing and/or is socially dispreferred.

Excerpt 5: Open class OIR (Albert et al., 2015)

- 1 SpA: We can't have three holidays in a year, can we?
- 2 SpB: Why not?
- 3 (1.5 seconds)
- 4 SpA: **Sorry?**
- 5 SpB: Why not?

These criteria lead us to focus on other-initiations of repair (OIR; Schegloff et al., 1977). An OIR occurs when a listener diagnoses their own problem of hearing or understanding and chooses to signal that problem to their interlocutor (e.g., “Sorry?” in Excerpt 5). OIRs meet all criteria for this study. First, turns marked by OIRs induce extra cognitive processing. The interlocutor identifies, diagnoses, and attempts to cognitively ‘repair’ the problem on their own (Fodor and Inoue, 2000) before planning their response (an OIR). *Postdiction* is an example of one “extra” cognitive process that listeners may employ when comprehending trouble source turns (Gwilliams et al., 2018). When a listener processes an ambiguous speech sound, they maintain that speech sound in working memory. As the turn unfolds, the listener not only comprehends new information but also relates that information to the ambiguous speech sound, in hopes of disambiguating the ambiguity. Maintaining ambiguous information in working memory requires cognitive resources (Baddeley, 2003). Comparing new and old information requires extra cognitive processing. If the information is not disambiguated, the listener may produce an OIR. See Section 3.2 on Cognitive Variables for the cognitive variables examined in this study.

Second, OIRs are socially dispreferred. OIRs slow conversation down (threaten progressivity; Schegloff, 1979). In Excerpt 5, line 5 (“Why not?”) could have been answered right away. Instead, Speaker A initiated repair, delaying progressivity for at least two turns. In addition, OIRs suggest that the OIR producer could not understand, or was not paying attention to, their interlocutor (threaten face; Schegloff et al., 1977). See Sections 3.3 and 3.4 for the social variables examined in this study.

Finally, OIRs are delayed in comparison to most other turns (Kendrick, 2015b; Schegloff et al., 1977). While responses to yes/no questions follow a median delay of approximately 340ms, OIRs follow a median delay of approximately 721ms (Kendrick, 2015b). These characteristics — increased cognitive processing, dispreferred status, and delay — make OIRs a good candidate for studying the relative influence of cognitive and/or social factors on utterance timing in conversation. Few, if any, other types of utterances have been shown in the literature to meet these criteria.

2. Method

2.1 Data Source

We collected OIRs from the Conversation Analytic British National Corpus (CABNC; Albert et al., 2015). The CABNC is a collection of the spoken language in the British National Corpus (BNC; Leech, 1992). The CABNC contains approximately 4.2 million words from 1,436 conversations recorded in the UK in the late 20th century. The participants were volunteers selected from different age groups, regions, and social classes in a demographically balanced way. We analyzed CABNC recordings of in-person (i.e., not over the telephone) and non-institutional (e.g., no call-ins to radio shows) conversations because these represent the most informal types of conversation. Exploratory, hypothesis-generating studies on conversation should analyze as unconstrained data as possible.

2.2 OIR Collection

Even within the category of OIRs, there were potential confounds. First, it takes more time to initiate longer utterances (Ferreira, 1991; Kempen and Huijbers, 1983; Sternberg et al., 1978). Therefore, we restricted our analysis to single-word OIRs. In addition, some OIRs perform multiple actions. For example, OIRs that repeat the trouble source turn can also challenge the relevance of that turn (Kendrick, 2015a; Robinson and Kevoe-Feldman, 2010). To complicate matters, some turns may either be an OIR or a different action, such as a joke (e.g., Excerpt 6). We included a potential OIR only if the interlocutors treated the potential trouble source as either misheard or misunderstood (Sacks et al., 1978). The interlocutors in Excerpt 6 treat “the concrete jungle” as a joke. Therefore, we would not include Excerpt 6 in our OIR collection.

Excerpt 6: an OIR as a joke, from Kendrick (2015a)

1 Cha: (It's a) nice place to work though.
2 (0.9)
3 **Liz: Ehhh what.=the concrete jungle,**
4 (0.2)
5 Cha: Aww:::::.=I think it's quite pretty. It has
6 ree:ds.

These constraints restricted our analysis to a small subset of all OIRs. Only approximately 25% of the OIRs in Kendrick (2015a) would have been included in our collection. However, we required a high level of precision to isolate the effects of specific cognitive and social variables in naturally occurring data.

Speakers can produce OIRs in the turn directly after the trouble source turn, or some number of turns removed from the trouble source turn. In this paper, we only analyzed OIRs that occur in the turn after the trouble source turn. This constraint made OIR FTO a more valid measure of the listeners' processing of the trouble source turn. In other sequences, where the OIR is produced later in the conversation, operationalization of some measures becomes more difficult. In Excerpt 7, Speaker 2 produces a joke on line 2 (evidenced by “it was a joke” on line 8). This joke makes laughter relevant in the next turn. Instead of laughing, Speaker 1 responds with “yeah” (line 3), a continuer or backchannel (Yngve, 1970; Schegloff, 1982) that implies that Speaker 1 is listening to Speaker 2. The mismatch prompts Speaker 2 to upgrade their original joke with “every weekend yeah” (line 5). Only after this upgrade does Speaker 1 produce an OIR. In this excerpt, should the

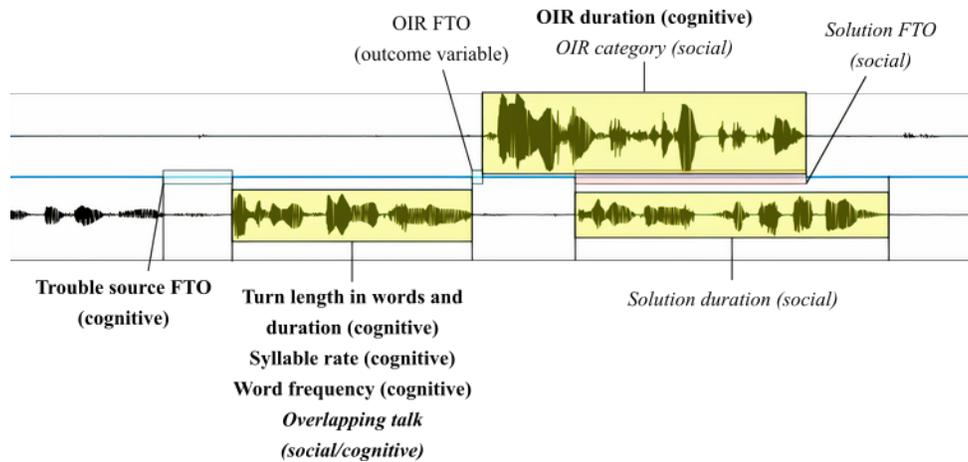


Figure 1: Measures in this study. Cognitive measures are bold and social measures are italicized.

trouble source include line 2, even though Speaker 1 responded to line 2 as if they did not notice any miscommunication? On one hand, Speaker 1 may not have noticed a communication problem until line 5. On the other hand, “every weekend” is not the entire joke targeted by the repair solution on line 8. To avoid these problematic methodological questions, we decided to only collect OIRs that occur in the turn directly after the trouble source. These selection criteria resulted in a final collection of 456 OIRs.

Excerpt 7: delayed OIR, collected by the Human Interaction Laboratory at Tufts University

- 1 SP1: I've never done some shit like this.
 2 SP2: yeah u- (0.3) oh I have li[ke,]
 3 SP1: [ye]ah,
 4 (0.3)
 5 SP2: every weekend yeah,
 6 **SP1: w:hat,**
 7 (4.1)
 8 SP2: I was just kidding. it was a joke,

3. Measures

In this section, we describe Floor Transfer Offset (FTO; de Ruiter et al., 2006), a quantification of the time between turns. Then, we describe the cognitive variables analyzed in this study. Next, we describe the social variables analyzed in the study. Finally, we describe our operationalization of overlapping talk during the trouble source. Figure 1 displays the measures analyzed in this study.

3.1 Floor Transfer Offset

Floor Transfer Offset (FTO; de Ruiter et al., 2006) is the time between two turns, calculated by subtracting the end time of the first turn from the start time of the second turn. Negative FTOs indicate overlaps, positive FTOs indicate gaps, and FTOs of 0 indicate gapless transitions.

We defined turn boundaries based on Turn Construction Units (TCUs) and Transition Relevance Places (TRPs; Clayman (2013); Sacks et al. (1978)). TCUs are composed of speech that is semantically, pragmatically, and prosodically complete. TRPs are moments where the preceding speech may be a TCU. At a TRP, the speaker may choose to end their TCU (and turn). In Figure 2, TRPs are marked with stars and the ends of TCUs are marked with circles. Line 2 ends in a TRP, but Speaker 1 adds two increments in lines 5 and 7.

```

1 *SP2:   Emma 「 Watson 1☆●
2 *SP1:   「 Emma 」 Watson has done a better 「 job breaking out 1☆
3 *SP2:   「 I know yeah☆● 」
4         (0.8)
5 *SP1:   but (.) I feel like Daniel Radcliffe just gave up☆
6 *SP2:   mm ☆●
7 *SP1:   and started doing weird movies☆●
    
```

Figure 2: Transcript with TRPs (stars) and TCUs (circles).

We calculated FTOs based on TRPs. We identified the TRP most relevant to the second turn. We subtracted the TRP time from the start time of the second turn. When FTOs are negative (indicating an overlap), the TRP is after the start point of the second turn. When FTOs are positive (indicating a gap), the TRP was before the start time of the second turn. The transcript displayed in Figure 2 would result in the FTOs displayed in Figure 3.

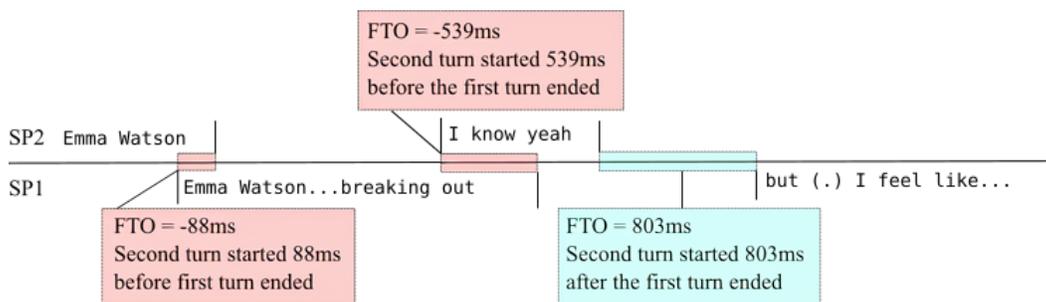


Figure 3: FTOs from lines 1-5 in Figure 2. Negative FTOs are represented in red, and the positive FTO is represented in blue.

We defined the trouble source turn as a single TCU that contained the problem. However, this presented some difficulties. For example, the trouble may arise from disordered self-repair across multiple TCUs. In Excerpt 8, lines 1-2 contain 3 short TCUs: “Sam’s dead died” “Dot died” and “His wife died in childbirth.” This sequence of self-repairs may have caused the communication problem. Excerpt 8 also demonstrates another problem: anaphora. When Ivy says “his” wife, she is referring to “Sam’s” wife. This reference would be unclear without the context of the immediately prior turn.

Excerpt 8: OIR sequence with a duration ratio of 0.06 from the CABNC

```

1      Ivy: Sam's dead die:d. Dot die:d his wife (1.3) died
2          in childbirth.
    
```

3 (0.9)
 4 **Con: Who.**
 5 (0.3)
 6 Ivy: Dot,
 7 (0.7)
 8 Con: Oh Sam's wife,

A third problem is when the trouble source turn was followed by a second TCU by the same speaker. In Excerpt 9, the OIR eventually targets “over there.” However, Gordon produces another TCU: “Is that it?” In each of these cases, we defined the trouble source turn as encompassing more than one TCU. There were 21 such cases. We conducted the statistical analyses with and without these sequences and obtained similar results.

Excerpt 9: OIR sequence the trouble source in the first TCU from the CABNC

1 GOR: What's that l- thing over there. Is that it?
 2 (0.4)
 3 DEB: **Whe:re.** (0.2) No that's Sean's and Kirsty's.

The FTO between the trouble source and the OIR (the OIR FTO) was our outcome variable. The FTO before the trouble source (the trouble source FTO) was our measure of discourse context: long trouble source FTOs indicate less discourse context (see Section 3.2 on Cognitive Variables). Finally, the FTO between the OIR and the repair solution (the repair solution FTO) was one measure of progressivity: the shorter the repair solution FTO, the less the OIR paused progressivity (see Section 3.3 on Social Variables). We marked all time points in Praat (Boersma, 2006).

3.2 Cognitive Variables

We measured the following cognitive variables: (1) duration (e.g., Piquado et al., 2010), (2) word frequency (for a review, see Goldinger, 1996), (3) syllable rate (Osada, 2004; Roberts et al., 2015) and (4) availability of discourse context for the trouble source. In this section, we operationalize and present hypotheses about each of these variables.

We measured trouble source turn length in duration and number of syllables. These variables were related, but not identical. Trouble source turns with more syllables tended to be longer, but a turn composed of many fast syllables could have the same duration as one composed of a few slow syllables. In addition, when the speaker paused mid-turn, utterance duration increased while the number of syllables remained the same. We predicted that longer trouble source turns would contain more information and would require more cognitive processing to comprehend (Roberts et al., 2015) As a result, we hypothesized that longer trouble source turns would be followed by longer OIR FTOs.

We measured the number of syllables in the trouble source turn and calculated the trouble source turn syllable rate. Roberts et al. (2015) used average phone lengths to calculate deviance from expected syllable rates. In contrast, we calculated the raw syllable rate. Trouble source turn produced quickly should have a higher information density on average. Therefore, faster trouble source turns should be associated with longer OIR FTOs. Fourth, we calculated the average word frequency of the trouble source from the SUBTLEX-UK corpus (van Heuven et al., 2014). Frequent words are easier to process (Goldinger, 1996). We predicted that the higher the average word frequency, the shorter the OIR FTO.

The relationship between discourse context and cognitive processing is more complex. In this paper, we use the term “discourse context” to refer to the sequential environment that occurs immediately prior to the trouble source turn. Discourse context constrains the upcoming turn. For example, the turn after a question is expected to be an answer (Schegloff, 1968). Listeners use this information to predict the upcoming speech act (Levinson, 2017; de Ruiter and Cummins, 2012; Gisladdottir et al., 2015). These predictions are time dependent: the more a turn is delayed (has a larger FTO), the more likely it is that the turn will be dispreferred (Kendrick and Torreira, 2015). However, dispreferred utterances are typically indirect and may take many forms, in contrast to preferred utterances which are typically direct (Pomerantz, 1984). Therefore, it may be more difficult to predict the exact words in a delayed trouble source turn (Gisladdottir et al., 2015).

Prediction may either increase or decrease the processing needed to notice a hearing or understanding problem. If a listener has a strong prediction for the trouble source turn, they may easily notice deviations from that prediction. In contrast, surprised listeners may work harder to make sense of the surprising information. This extra cognitive processing would delay the OIR.

Discourse context and listener predictions about the trouble source should have the same effect on OIR FTO, because discourse context leads to prediction. Therefore, we can make bidirectional predictions about the effect of discourse context on OIR FTOs. We operationalized discourse context as the FTO before the trouble source turn; the shorter the trouble source turn FTO, the stronger the context for the trouble source turn. We also created a binary variable indicating whether the trouble source turn FTO was greater than 1.5 seconds. Most turns transitions are shorter than 1 second, so trouble source turns with FTOs longer than 1.5 seconds are considered to be *disjointed*, or with little discourse context.

3.3 Social Variables

In this paper, we focused on two preferences especially relevant to the timing of OIRs: progressivity and specificity. Speakers aim to be progressive, or to continue the activity of the conversation (Schegloff, 2007; Stivers and Robinson, 2006). OIRs stop the current activity for at least two turns: the OIR and the repair solution. OIRs always decrease progressivity, but some require more time and work to resolve than others. We predict that speakers initiate OIRs earlier when they require less time to resolve.

We measured repair solution FTO by subtracting the end time of the OIR from the start time of the repair solution. In addition, we calculated the duration of the repair solution. However, the duration of the repair solution is dependent on the duration of the trouble source turn. Even when an OIR is more progressive than another would have been, if the trouble source turn is very long, the repair solution will also be very long. Therefore, we calculated the *duration ratio* of the sequence: the ratio of the repair solution duration to the trouble source turn duration. A duration ratio of 1 signifies that the repair solution and trouble source turn were the same length. A duration ratio of 0.5 signifies that the repair solution was half the length of the trouble source turn. The sequence in Excerpt 8 has a very small duration ratio, in part because Ivy pauses for 1.3 seconds during the trouble source turn. When solution FTO was longer or duration ratio was higher, progressivity was paused for longer. We predicted that in such sequences, OIR FTO would also be longer.

Dingemanse et al. (2015) examined measures that were similar but distinct from the ones explored in this study. First, they found that the duration of the trouble source turn was always approximately equal to the combined duration of the OIR and repair solution. At first glance, this

evidence suggests that our measure of duration ratio would be around 1 for all sequences. However, Dingemanse et al. (2015) did not focus on single-word OIRs. The variance in duration of single-word OIRs is much lower than the variance in duration of all OIRs. Different single-word OIRs still request different types of repair solutions, so we still expected variance in the repair solution duration. We also did not include the duration of the OIR in our duration ratio measure. Therefore, we still expected variance in duration ratios.

In addition, there is a preference for more specific OIRs (Schegloff et al., 1977; Dingemanse et al., 2015; Kendrick, 2015b). Specificity is related to the size of the problem marked by the OIR; specific OIRs mark “smaller” communication problems (e.g., Excerpts 8 and 9). Open class OIRs (e.g., Excerpts 5 and 7) mark the entire trouble source turn (Drew, 1997). There are two reasons that specific OIRs may be preferred over open class OIRs. First, it may be more face-threatening to admit to a larger communication problem by producing an open class OIR. Second, since specific OIRs mark smaller problems than open class OIRs, the repair solutions to specific OIRs tend to be shorter (Clark and Schaefer, 1987; Dingemanse et al., 2015; Grice, 1975). This means that specific OIRs halt progressivity for a shorter duration.

We created a binary variable indicating whether the OIR was open class. Out of the 456 OIRs, 269 (59%) were open class OIRs and 187 (41%) were specific OIRs. For every specific OIR, we had almost 1.5 open class OIRs. In contrast, Kendrick (2015b) found 2.2 specific OIRs for every open class OIR. We believe this difference is due to our collection criteria. Many repeats, candidate understandings, restricted requests, and other types of specific OIRs tend to have more than one word and, therefore, were not included in this collection. We predicted that specific OIRs would be initiated earlier than open OIRs, the repair solutions in specific OIR sequences would be shorter, and that duration ratio will contribute to earlier specific OIRs.

Parts of the interlocutors’ social identity could have effects on OIR FTO. For example, the relationship between gender and turn-timing has been a controversial topic for decades (e.g., West and Zimmerman, 2015). Roberts et al., (2015) found that for each male in a dyad, FTOs increased by approximately 70ms. However, we decided to not analyze social identity in this analysis. First, the CABNC has inconsistent labels of these variables. Second, even if we were to find that members of one group (e.g., women) initiate OIRs earlier than members of another group (e.g., men), we would not be able to explain why. Most likely, differences between social groups would be due to a complex underlying variable, such as dominance. Third, identity is not easily categorized. For example, even if we limited the investigation of gender to a binary conceptualization, and only to cisgender individuals, the categories of “cisgender male” and “cisgender female” are still very broad. The stereotypes associated with White and Black women are very different (Landrine, 1985). Even within one of those subcategories, different groups of White or Black women are perceived as more or less feminine and are expected to conform to stereotypes to different degrees. For example, female athletes are often perceived as less feminine (Kauer and Krane, 2006). The relationship between social identity and turn-taking is interesting and a worthwhile subject, but requires more detailed analysis than was possible in this paper.

3.4 Overlapping Talk

Finally, we analyzed overlapping talk during the trouble source turn. We hypothesized that overlapping talk could have either social or cognitive effects on OIR FTO. Background speech interferes with comprehension, even more than meaningless noise (Koelewijn et al., 2015, 2012; Schneider

et al., 2007). This finding suggests that listeners extract and process the information contained in both the target speech (from their interlocutor) and overlapping speech (from the background). If so, overlapping talk during the trouble source turn should increase cognitive processing, resulting in delayed OIRs.

Overlapping talk during the trouble source turn may also have a social effect on OIR FTOs. Participants work to maintain a positive public image, or *face* (Goffman, 1955; Lerner, 1996). When an utterance is face-threatening, that utterance is dispreferred and delayed. OIRs are potentially face-threatening because they could be attributed to poor language comprehension, outsider status, or a lack of attention. If someone talks over the trouble source turn, the OIR can be attributed to something other than the OIR speaker. In such cases, the OIR should be less face threatening. In contrast to the cognitive hypothesis, the social hypothesis predicts that overlapping talk during the trouble source turn should be related to shorter gaps before the OIR.

We operationalized overlapping talk as speech produced either by an interlocutor within the interaction, people in the background, or a television or radio, before the start of the OIR. We did not analyze music or non-speech noises. While environmental noise has been linked to more open class OIRs (Dingemanse et al., 2015), given the data collection methods, noises that were loud to the transcriber may not have been loud to the interlocutors. We calculated two values: a binary variable indicating whether there was any overlapping speech during the trouble source turn, and the percentage of the trouble source turn that was overlapped.

4. Analysis

In natural conversation, it is impossible to isolate the effect of one predictor from all the others. Some statistical tests, such as principal components analyses or decision trees, allow researchers to simulate several effects at once. However, many of these tests provide output that is difficult to interpret. Multiple regression solves both problems. Multiple regression can analyze multiple effects at once, while calculating the effect of each predictor on its own.

The commonly used Frequentist approach to inferential statistics computes the probability of the data or more extreme data under a null hypothesis, the alternative hypothesis being the negation of this null hypothesis. In contrast, Bayesian statistics allow researchers to directly compare the predictive adequacy of different hypotheses. Given that our goal was to compare cognitive and social models of OIR FTO, we decided to use Bayesian multiple regression. In this paper, we will report Bayes Factors (BF_{ab}) representing the relative probability of the observed data under model *a* over model *b*. A larger Bayes Factor means that there is a larger difference between the two models. Typically, researchers report BF_{10} , and the larger it is, the more evidence there is for the alternative over the null (Kass and Raftery, 1995; Jeffreys, 1998); however, in this paper, we will also discuss evidence for cognitive over social factors.

We created three types of regression models. First, we analyzed models with only cognitive predictors. Second, we analyzed models with only social predictors. We compared the cognitive and social models directly to see which model explained more variance in OIR FTO. Finally, we created models that combined both social and cognitive predictors. For details on the regression models and residuals, see <https://osf.io/stna6/>. We performed all analyses with JASP (Team, 2018; jasp-stats.org). For access to the datafile, see <https://osf.io/rkuav/>.

4.1 Pre-screening

Regression models assume that the model residuals are normally distributed. Our outcome variable, OIR FTO, was positively skewed, which resulted in positively skewed model residuals. To make the models more valid, we log-transformed the OIR FTOs. We added a value to each such that the smallest OIR FTO was 1, and then took the 10-based logarithm of the resulting value. As a result, the model does not assume a linear relationship between the predictors and OIR FTO, but between the predictors and the log-transformed OIR FTO. This means the effect of any independent variable on raw OIR FTO increased in magnitude as its value increased. As a result, we present effects as percentages – the percent difference in OIR FTO when a predictor variable is different. The percent difference indicates a greater raw difference when the value of the predictor variable is larger. The residuals of the resulting models were approximately normally distributed (see the annotated JASP file at <https://osf.io/stna6/> for regression residuals).

In addition, regression models assume independence of predictor variables. When predictor variables are related, regression models have greater false negative rates and inaccurately estimated effects. We found strong evidence for several relationships between predictors (for details on our prescreening, see the annotated JASP file at <https://osf.io/fznht/>). To mitigate the consequences of these relationships, we used JASP to compare all combinations of predictors in each category within regression model (cognitive, social, and overall). JASP determined the most predictive model within each category. Below, we report the models that best predict log OIR FTO.

Finally, we determined whether overlapping talk during the trouble source turn should be added to the cognitive or social regression model. A Mann-Whitney U test found strong evidence that overlapped trouble source turns (Median = 478ms) were followed by shorter OIR FTOs than non-overlapped trouble source turns (Median = 596ms, $BF_{10} = 10.37$). Further, a Kendall correlation found decisive evidence that OIR FTO was shorter when a greater proportion of the trouble source turn was overlapped ($t = -0.13$, 95% Bayesian Credible Interval, henceforth referred to as CI = $-0.19 - -0.07$, $BF_{10} = 421.93$). This evidence suggested that overlapping talk during the trouble source turn may make OIRs less socially dispreferred. Therefore, we included overlapping talk during the trouble source turn in the social regression model.

4.2 Missing Data

Out of 456 OIRs, 307 (67.32%) had data for all predictor variables. The relatively high rate of missing data is a common consequence of collecting natural data: we could not control for environmental noise or ask participants to speak into the microphone. This means that some data were excluded from the regression models. In Excerpt 10, for example, Speaker B resolved the problem on their own so there was no value for repair solution FTO or duration. For each statistically significant effect found in the final models below, we performed a simple analysis including only the relevant predictor. These simple tests allowed us to analyze all the OIRs with values for OIR FTO and the relevant predictor, even if there was missing data for another predictor. This also allowed us to use nonparametric versions of tests on raw OIR FTO. For example, we conducted a Mann-Whitney U test to investigate the effect of OIR category (specific vs. ‘open’ class) on OIR FTO. Any OIR with a value for OIR category was included in this test.

Excerpt 10: Resolving OIR without verbal repair solution

1 Spa: Nice little bu:m there,
2 (1.0)

3 **Spb: Where.**
 4 (0.5)
 5 Spb: (laughing) Oh yeah.

To ensure that there were no relationships between missing data and OIR FTO, we created a binary variable for each predictor indicating whether data were missing or present. We then performed a series of Bayesian Mann-Whitney U-tests to determine whether OIR FTO differed when each predictor was absent from the model. We found no evidence that OIR FTO differed when any predictor variable was missing. For details on missing data, see our JASP file at <https://osf.io/h5ysb/>.

4.3 Cognitive Factor Regression

We regressed the log-transformed OIR FTO on all cognitive factors (trouble source turn FTO, disjointed status, length in both words and duration, average word frequency and syllable rate). The most predictive model included trouble source turn duration and FTO, and the binary variable indicating that the trouble source turn was disjointed ($R^2 = 0.08$, $BF_{10} = 20.69$). See Table 1 for the unstandardized coefficient results for cognitive factors included in the best cognitive model.

Predictor	Mean	Predictor Std. Deviation	95% CI	$BF_{inclusion}$
TS Duration	-2.90%	1.13 seconds	-4.60% – -1.10%	1.47
TS FTO	0.40%	6.81 seconds	0.05% – 0.70%	3.93
TS = Disjointed	5.10%	–	0.8% – 9.50%	4.66

Table 1: Regression Coefficients for Cognitive Variables.

4.4 Social Factor Regression

We regressed the log-transformed OIR FTO on all social factors (the percent overlap in the trouble source turn, the binary variable indicating any overlap during the trouble source turn, OIR category and duration ratio). The most predictive model consisted of OIR category and the binary variable indicating overlap in the trouble source turn ($R^2 = 0.04$, $BF_{10} = 11.41$). We found anecdotal evidence (the classification “anecdotal” is taken from Appendix B in Jeffreys, 1998) that when the trouble source turn was overlapped, the OIR FTO was 5.30% shorter (95%CI = -9.20% — -1.10%, $BF_{inclusion} = 3.11$). We found substantial evidence that when the OIR was open class, the OIR FTO was 6.00% longer (95%CI = 2.10% — 9.80%, $BF_{inclusion} = 8.11$).

4.5 Cognitive and Social Model Comparison

Next, we compared the social and cognitive models in three stages. First, we added all cognitive factors to the most predictive social model (OIR category and the binary variable indicating overlap in the trouble source turn). The most predictive model added trouble source turn FTO and trouble source turn duration to the most predictive social model ($BF_{10} = 243.39$, $R^2 = 0.10$). See Table 2 for the unstandardized coefficients for the resulting model.

Second, we added all social factors to the most predictive cognitive model (trouble source turn duration, trouble source turn FTO, and the binary variable indicating disjointed trouble source turns). The most predictive model added all social factors: OIR category, duration ratio, the percent

	Predictor	Mean	Predictor Std. Deviation	95% CI	BF _{inclusion}
Cognitive	TS Duration	-2.60%	1.13 seconds	-4.40% – -0.80%	1.66
	TS FTO	0.40%	6.81 seconds	0.20% – 0.70%	4.34
Social	OIR = Open Class	4.10%	–	0.30% – 7.80%	1.00
	TS = Overlapped	-6.00%	–	-10.00% – -2.00%	1.00

Table 2: Regression Coefficient Table for Cognitive Variables Added to Social Model.

overlap in the trouble source turn, and the binary variable indicating overlap in the trouble source turn ($BF_{10} = 67.30$, $R^2 = 0.11$). See Table 3 for the unstandardized results for this model.

Finally, we compared the models directly to determine whether one model was more likely than the other. We found anecdotal evidence that the cognitive model was more predictive than the social model ($BF_{10} = 1.94$).

	Predictor	Mean	Predictor Std. Deviation	95% CI	BF _{inclusion}
Cognitive	TS Duration	4.00%	1.13 seconds	0.40% – 7.70%	1.00
	TS FTO	1.30%	6.81 seconds	0.70% – 1.90%	1.00
	TS = Disjointed	-3.60%	–	-13.30% – 6.00%	1.00
Social	OIR = Open Class	-14.30%	–	-22.70% – -6.00%	67.58
	Duration Ratio	6.50%	0.88	1.80% – 11.30%	13.83
	% of TS Overlapped	-0.10%	25.64%	-0.30% – 0.10%	1.49
	TS = Overlapped	-1.00%	–	-14.80% – 12.70%	0.97

Table 3: Regression Coefficient Table for Social Variables Added to Cognitive Model.

4.6 Overall Model

We regressed the log-transformed OIR FTO on all predictor variables. The most predictive model included both cognitive (trouble source turn duration, trouble source turn FTO) and social (the binary variable marking overlapped trouble source turns, OIR category) variables ($R^2 = 0.11$, $BF_{10} = 33848.68$). We found strong evidence for the overall model over the social model ($BF_{10} = 12.36$) and substantial evidence for the overall model over the cognitive model ($BF_{10} = 6.35$). See Table 4 for the unstandardized coefficients. The residuals approximated a normal distribution.

	Predictor	Mean	Predictor Std. Deviation	95% CI	BF _{inclusion}
Cognitive	TS Duration	-3.00%	1.13 seconds	-5.00% – -1.00%	1.14
	TS FTO	0.40%	6.81 seconds	0.10% – 0.70%	2.85
Social	OIR = Open Class	5.20%	–	1.00% – 9.40%	2.47
	TS = Overlapped	-5.50%	–	-9.80% – -1.10%	1.10

Table 4: Coefficient Table for Overall Model (Unstandardized).

5. Discussion

Interlocutors initiated OIRs later when the trouble source turn (1) had less discourse context, (2) was not overlapped by other talk, or (3) was shorter. Finally, open class OIRs were initiated later than specific OIRs. Adding cognitive variables to the social model, or social variables to the cognitive model, improved model predictions. The best overall model included both types of measures.

5.1 Discourse Context

When the trouble source turn FTO was longer, OIR FTO was longer. 4 shows that this effect persisted in a simple post-test ($t = 0.15$, $BF_{10} = 538.16$). During the pre-screening process, we also found substantial evidence that a greater proportion of disjointed trouble source turns ($\approx 67\%$) were more followed by open class OIRs than non-disjointed trouble source turns ($\approx 54\%$).

Listeners use discourse context to anticipate features of the trouble source turn. Accurate predictions decrease the information the listener has to process, either by pre-activating upcoming representations or easing the integration of new information (Kuperberg and Jaeger, 2016). Therefore, it may be that discourse context helps the listener process the correctly predicted information in the trouble source turn, reducing the time needed to respond. In addition, if a communication problem violates a prediction, it may be easier to notice. Neurological evidence suggests that listeners process speech act mispredictions approximately 300ms after the start of the turn (Gisladottir et al., 2015). This is just a few words into the unfolding turn. At the same time, violated speech act predictions have been related to more open class OIRs (Drew, 1997), suggesting that prediction can increase the size of the comprehension problem. Together, these findings suggest that predictions help the listener identify problems, but make it harder for the listener to diagnose or resolve problems.

Attention may also contribute to this effect. Some of our trouble source turns had long FTOs, even over 30 seconds. These trouble source turns likely started a sequence. From the audio recordings, it appeared that many occurred while the interlocutors were engaged in a concurrent activity, such as walking down the street or doing the dishes. However, we do not have access to the video data to confirm whether the listener was distracted or multitasking. In addition, no trouble source turns began with greetings (e.g., “Hi,”) or address terms (e.g., “Barbara, what are you up to?”). Such prefaces attract the listener’s attention. The listeners may not have been attending to trouble source turns with very long FTOs. Listeners still process the general characteristics, but not the content, of speech they are not currently attending to (Arons, 1992). This hypothesis may explain

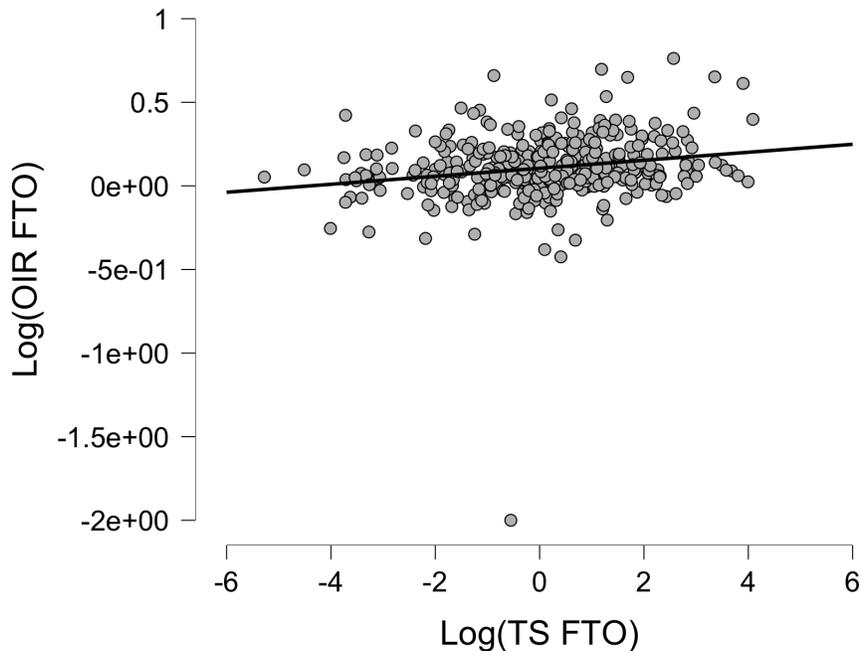


Figure 4: Positive correlation between $\log(\text{trouble source FTO})$ and $\log(\text{OIR FTO})$

why interlocutors produce more open OIRs when multitasking (Dingemanse et al., 2015). Particularly intriguing is the finding that a listener requires approximately 170ms to switch their attention to, and begin to process the content of, an utterance (Cherry and Taylor, 1954). This is the exact difference in the median OIR FTO after disjointed (Median = 674ms) and non-disjointed (Median = 504ms) trouble source turns. This evidence raises several new and intriguing questions about the role of attention in conversation. Do listeners respond more quickly to disjointed turns that start with address terms? Are all disjointed utterances harder to process, or is this effect unique to trouble source turns that are especially difficult to comprehend? In summary, one explanation for the increased OIR FTO after trouble source turns with less discourse context is that the listener needed more time to direct their attention to the trouble source turn.

Recent work in conversation analysis suggests that there are different types of lapses in conversation (Hoey, 2015). Some lapses are expected given the previous conversation, some are accounted for by parallel activities, and some are conspicuous. In this paper, we did not distinguish between these lapses. Future research could examine whether the type of lapse preceding the trouble source turn influences the timing of OIRs. One possibility is that conspicuous lapses draw the interlocutors' attention, so when a response is finally given it will be easier to process. In contrast, if a gap follows the end of a sequence, interlocutors may find it more difficult to process the utterance after this gap.

Another alternative explanation could be that all FTOs in a conversation are correlated; conversations are either fast or slow. However, repair solution FTO was not correlated with OIR FTO ($t = 0.08$, $\text{BF}_{10} = 0.93$, 95% CI = 0.15 – 0.01), although repair solution FTO was positively correlated with trouble source turn FTO ($t = 0.11$, $\text{BF}_{10} = 15.65$, 95% CI = 0.18 – 0.05). It is difficult to

interpret these findings; future research should compare OIR FTOs to a distribution of FTOs from across the entire conversation.

5.2 Overlapping Talk During the Trouble Source Turn

Overlapped trouble source turns were followed by shorter OIR FTOs (Figure 5). This effect is discrete and not continuous: more overlapping talk does not relate to even shorter OIR FTOs. Overlapping talk during the trouble source turn could provide a face-saving explanation for the OIR; the interlocutors may attribute the problem to the environment, and not the person who produced the trouble source turn (Kelley, 1973). As a result, there is less social cost for OIRs after overlapped trouble source turns, regardless of the duration of the overlap. The OIR producer may not have to account for the OIR to the same degree (Robinson, 2016).

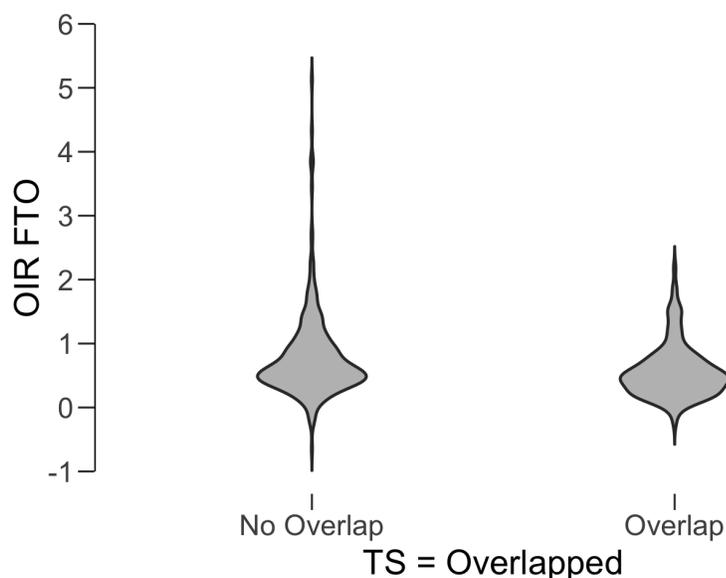


Figure 5: OIR FTO after overlapped and non-overlapped trouble source turns.

This result is the opposite of what we would expect if overlapping talk had a cognitive effect on timing in conversation. If listeners processed both target and overlapping speech, they would take longer to respond to an overlapped turn. The findings of this study do not rule out the possibility that overlapped turns take more cognitive processing to comprehend. However, listeners may not have the social prerogative to attempt to cognitively repair the problem if the OIR is not face threatening. Even if the interlocutor has not finished processing all the possible information from the turn, they may initiate repair if there is little cost for doing so.

Interestingly, during our data pre-screening we found that OIR category was not predicted by the presence of overlapping talk during the trouble source turn ($BF_{10} = 0.29$). This was unexpected — previous analyses found that open class repair initiations were more likely in loud or distracting environments (Dingemans et al., 2015). However, we did not analyze non-speech noises or distractions. Further research is needed to determine whether speech, other noises, and concurrent activities have different effects on OIRs.

We also did not distinguish between different sources of overlapping talk. If the trouble source turn overlapped with the previous turn (negative trouble source turn FTO), then the trouble source turn speaker may be considered responsible for the problem. In contrast, if the eventual OIR producer interrupted the trouble source, the responsibility may lie more with the OIR producer. Future research should explore the social and cognitive differences between general background noise, a negative trouble source FTO, and mid-turn overlap during the trouble source.

5.3 OIR Category

Open class OIRs (e.g., “Sorry?”) were initiated later than specific OIRs (e.g., “Who?”). This finding supports previous analyses of OIRs (Kendrick, 2015b; Schegloff et al., 1977) and preference organization (Roberts et al., 2015; Kendrick and Torreira, 2015). We extended these findings by accounting for other effects that affect timing in conversation.

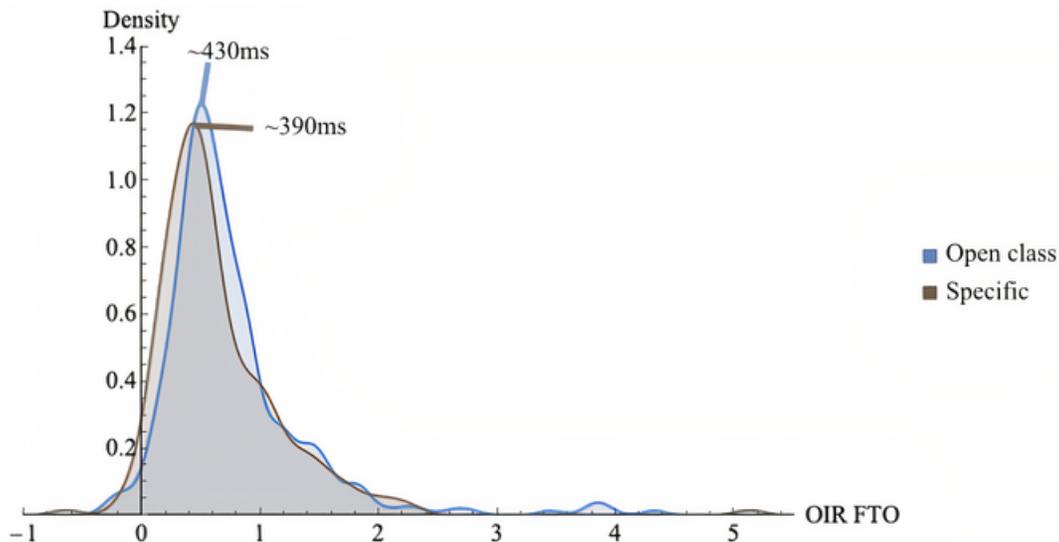


Figure 6: Probability density of OIR FTO separated by OIR category (open vs. specific)

We also attempted to understand why specific OIRs were produced after shorter OIRs. We hypothesized that interlocutors delay open class OIRs because open class OIRs are less progressive (Dingemanse et al., 2015). A Kendall’s correlation did find substantial evidence that duration ratio was slightly positively correlated with OIR FTO ($t = 0.10$, $BF_{10} = 5.92$). However, once we accounted for other variables, duration ratio did not explain unique variability in OIR FTO. Therefore, we suggest that the relationship between duration ratio and OIR FTO is indirect. Specific OIRs were initiated earlier, and were followed by shorter repair solutions, but specific OIRs were not initiated earlier because they were followed by shorter solutions. From a psychological perspective, this finding is not surprising. For the OIR producer to account for the expected length of the repair solution, they would first have to estimate the length of the repair solution. But of course the OIR producer does not know the repair solution. Further, simulating the repair solution duration would be another cognitively demanding task in a time-sensitive scenario.

In our pre-screening, we found substantial evidence that open class OIRs ($M = 387\text{ms}$) were followed by shorter repair solution FTOs than specific OIRs ($M = 506\text{ms}$, $W = 13307.00$, $BF_{10} = 5.68$). This finding contradicted our assumption that open class OIRs would be followed by longer repair solution FTOs, and therefore be even less progressive. This surprising result may be due to the nature of open vs. specific repair solutions. Solutions to open class OIRs may be more likely to repeat the trouble source. In contrast, solutions to specific OIRs may be more likely to incorporate new information or words. As a consequence, specific repair solutions may require more planning. Repair solutions highlight an interesting tension between cognitive and social factors in conversation that should be explored in a separate, focused analysis.

So, why are open class OIRs initiated later than specific OIRs? An examination of Figure 6 suggests two effects. First, the open class OIR FTO mode (approximately 430ms) is somewhat longer than the specific OIR FTO mode (approximately 390ms). Second, most OIR FTOs longer than 2.5 seconds precede open class OIRs. The first effect — the difference in modes — may be a cognitive effect. If open class OIRs mark larger comprehension problems, then the OIR speaker may spend more time attempting to cognitively repair that comprehension problem. The second effect — the right tail of the distribution — may be a social effect. If open class OIRs are more likely to be extremely dispreferred, then the interlocutor who produces the open class OIR may wait an especially long time to produce the OIR.

5.4 Trouble Source Turn Duration

Trouble source turns with longer durations were followed by shorter OIR FTOs. When the trouble source turn was longer, the listener had more time to plan their upcoming OIR. In contrast, Roberts et al. (2015) found that longer turns were followed by longer FTOs. However, Roberts et al., 2015 analyzed all utterances in an entire corpus. We restricted our analysis to a very small proportion of all utterances to control for a variety of confounds. It may be that confounds are responsible for the effect in Roberts et al. (2015). Another possible explanation is that OIRs have a unique relationship with the duration of the previous turn. It is possible that once listeners realize there is a comprehension problem, they stop comprehending the remaining information in the turn. Instead, they immediately begin planning their OIR. If so, then the amount of information contained in a trouble source turn is less relevant to the timing of OIRs. Either way, we found that longer trouble source turns were related to shorter OIR FTOs.

6. Conclusion

This study found that (1) discourse context, (2) overlapping talk during the trouble source turn, (3) trouble source turn duration, and (4) OIR category play a role in the timing of other-initiations of repair. First, we found that discourse context allows the listener to predict the upcoming turn. These predictions facilitate processing information that matches the prediction, and noticing information that violates the prediction. Therefore we concluded that the relationship between discourse context and OIR FTO is a cognitive effect.

Second, overlapped trouble source turns were followed by shorter OIR FTOs. If overlapping talk had a cognitive effect on OIR FTO, we would expect that OIR FTO would be longer (Koelewijn et al., 2012, 2015; Schneider et al., 2007). In addition, we would expect a continuous effect, where OIR FTO would be longer when there was more overlapping talk during the trouble source turn. However, we found that any overlapping talk during the trouble source resulted in similarly short-

ened OIR FTOs. When there is overlapping talk, the interlocutors can attribute communication problems to that talk; the OIR speaker is not accountable for the problem.

Third, longer trouble source turns were followed by shorter OIR FTOs. We believe this is a cognitive effect: the longer the trouble source turn, the more time the OIR producer has to plan the OIR. Finally, our analysis cannot conclusively determine whether OIR category has an exclusively social or cognitive effect on OIR FTO. Our hypothesis – that open class OIRs are especially delayed because they delay progressivity more than specific OIRs – was not supported by our analysis. It may be that open class OIRs are, in fact, more dispreferred than specific OIRs because they admit to a larger communication problem, and therefore may be more face-threatening. On the other hand, this effect may be cognitive effect: that open class OIRs are more likely when the trouble source turn recipient has a larger problem of hearing or understanding, and therefore induce more cognitive processing. Further research would be needed to determine if OIR category should be categorized as a cognitive or social factor.

More generally, these findings show that both cognitive and social factors play a role in the timing of OIRs. Interlocutors are constrained by aspects of cognitive processing: they cannot speak until they are cognitively ready to do so. Cognitive limitations put a lower bound on the time required before initiating speech. However, interlocutors are also sensitive to social rules: they sometimes are cognitively able to initiate speech earlier, but nevertheless delay because of social factors. So the most important general finding from our study is that the timing of an utterance is influenced by both cognitive and social constraints.

Surprisingly, we found strong evidence for the null-hypothesis that there is no effect of average word frequency in the trouble source turn on OIR FTO ($BF_{01} = 14.76$). Many studies have found that listeners respond to frequent words more quickly (Goldinger, 1996). Perhaps this effect is only detectable using highly controlled experimental paradigms like lexical decision or picture naming. Another possibility is that interlocutors in conversation do not process language in the same way as participants in experiments. Or it may be that interlocutors process trouble source turns differently than other turns; Roberts et al. (2015) did find that word frequency predicted turn transition times in a corpus of natural data. It may be that trouble source turns are unique in that the frequency of the words in the turn is less influential. This finding raises a provocative question: how many well-established psycholinguistic findings, like frequency effects or semantic priming, are relevant in natural interaction, and if they are relevant, in which types of sequences?

Conversation, one of the crown jewels of human cognition, requires the streamlined interaction of a variety of skills and abilities. In this paper, we examined how social attribution, attention, attribution, and prediction come together when humans resolve miscommunication. To understand the mechanisms underlying conversational interaction we have to take into account both cognitive social constraints.

Acknowledgements

We would like to thank Dr. Saul Albert and Dr. Ariel Cohen-Goldberg for their guidance in the early stages of this project. Funding for this project was provided by the United States Air Force Laboratory.

References

- Saul Albert, Laura E. de Ruiter, and Jan P. de Ruiter. CABNC: The Jeffersonian transcription of the Spoken British National Corpus, 2015.
- Barry Arons. A review of the cocktail party effect. *Journal of the American Voice I/O Society*, 12: 35–50, 1992. doi: 10.1.1.30.7556.
- Alan Baddeley. Working memory and language: An overview. *Journal of communication disorders*, 36(3):189–208, 2003.
- Paul Boersma. Praat: Doing phonetics by computer. <http://www.praat.org/>, 2006.
- Sara Bögels, Kobin H. Kendrick, and Stephen C. Levinson. Conversational expectations get revised as response latencies unfold. *Language, Cognition and Neuroscience*, 35(6):766–779, 2020.
- E. Colin Cherry and W. K. Taylor. Some further experiments upon the recognition of speech, with one and with two ears. *The Journal of the acoustical society of America*, 26(4):554–559, 1954.
- Herbert H. Clark and Jean E. Fox Tree. Using uh and um in spontaneous speaking. *Cognition*, 84(1):73–111, 2002.
- Herbert H. Clark and Edward F. Schaefer. Collaborating on contributions to conversations. *Language and cognitive processes*, 2(1):19–41, 1987.
- Steven E. Clayman. Sequence and solidarity. In *Advances in group processes*. Emerald Group Publishing Limited, 2002.
- Steven E. Clayman. Turn-constructive units and the transition-relevance place. *The handbook of conversation analysis*, pages 150–166, 2013.
- Jan P. de Ruiter and Chris Cummins. A model of intentional communication: AIRBUS (Asymmetric Intention Recognition with Bayesian Updating of Signals). *Proceedings of SemDial 2012*, pages 149–50, 2012.
- Jan P. de Ruiter, Holger Mitterer, and Nick J. Enfield. Projecting the end of a speaker’s turn: A cognitive cornerstone of conversation. *Language*, 82(3):515–535, 2006.
- Mark Dingemanse, Seán G. Roberts, Julija Baranova, Joe Blythe, Paul Drew, Simeon Floyd, Rosa S. Gisladottir, Kobin H. Kendrick, Stephen C. Levinson, Elizabeth. Manrique, et al. Universal principles in the repair of communication problems. *PloS one*, 10(9):e0136100, 2015.
- Franciscus Cornelis Donders. On the speed of mental processes. *Acta psychologica*, 30:412–431, 1969.
- Paul Drew. ‘Open’ class repair initiators in response to sequential sources of troubles in conversation. *Journal of pragmatics*, 28(1):69–101, 1997.
- Fernanda Ferreira. Effects of length and syntactic complexity on initiation times for prepared utterances. *Journal of Memory and Language*, 30(2):210–233, 1991.

- Janet Dean Fodor and Atsu Inoue. Syntactic features in reanalysis: Positive and negative symptoms. *Journal of Psycholinguistic Research*, 29(1):25–36, 2000.
- Rosa S. Gisladottir, Dorothee J. Chwilla, and Stephen C. Levinson. Conversation electrified: ERP correlates of speech act recognition in underspecified utterances. *PloS one*, 10(3):e0120068, 2015.
- Erving Goffman. On face-work: An analysis of ritual elements in social interaction. *Psychiatry*, 18(3):213–231, 1955.
- Stephen D. Goldinger. Auditory lexical decision. *Language and Cognitive Processes*, 11(6):559–568, 1996.
- Herbert P. Grice. Logic and conversation. In *Speech acts*, pages 41–58. Brill, 1975.
- Laura Gwilliams, Tal Linzen, David Poeppel, and Alec Marantz. In spoken word recognition, the future predicts the past. *Journal of Neuroscience*, 38(35):7585–7599, 2018.
- Alexa Hepburn and Galina B. Bolden. The conversation analytic approach to transcription. *The handbook of conversation analysis*, 57:76, 2013.
- John Heritage. Well-prefaced turns in English conversation: A conversation analytic perspective. *Journal of Pragmatics*, 88:88–104, 2015.
- Elliott M. Hoey. Lapses: How people arrive at, and deal with, discontinuities in talk. *Research on Language and Social Interaction*, 48(4):430–453, 2015.
- Peter Indefrey and Willem J. M. Levelt. The spatial and temporal signatures of word production components. *Cognition*, 92(1-2):101–144, 2004.
- Stefanie Jansen, Hendrik Wesselmeier, Jan P. de Ruiter, and Horst M. Mueller. Using the readiness potential of button-press and verbal response within spoken language processing. *Journal of neuroscience methods*, 232:24–29, 2014.
- Harold Jeffreys. *The theory of probability*. OUP Oxford, 1998.
- Robert E. Kass and Adrian E. Raftery. Bayes factors. *Journal of the american statistical association*, 90(430):773–795, 1995.
- Kerrie J. Kauer and Vikki Krane. “Scary dykes” and “feminine queens”: Stereotypes and female collegiate athletes. *Women in Sport & Physical Activity Journal*, 15(1):42, 2006.
- Harold H. Kelley. The processes of causal attribution. *American psychologist*, 28(2):107, 1973.
- Gerard Kempen and Pieter Huijbers. The lexicalization process in sentence production and naming: Indirect election of words. *Cognition*, 14(2):185–209, 1983.
- Kobin H. Kendrick. Other-initiated repair in English. *Open Linguistics*, 1, 2015a.
- Kobin H. Kendrick. The intersection of turn-taking and repair: The timing of other-initiations of repair in conversation. *Frontiers in psychology*, 6(250):10–3389, 2015b.

- Kobin H. Kendrick and Francisco Torreira. The timing and construction of preference: A quantitative study. *Discourse Processes*, 52(4):255–289, 2015.
- Thomas Koelewijn, Adriana A. Zekveld, Joost M. Festen, and Sophia E. Kramer. Pupil dilation uncovers extra listening effort in the presence of a single-talker masker. *Ear and Hearing*, 33(2): 291–300, 2012.
- Thomas Koelewijn, Hilde de Kluiver, Barbara G. Shinn-Cunningham, Adriana A. Zekveld, and Sophia E. Kramer. The pupil response reveals increased listening effort when it is difficult to focus attention. *Hearing research*, 323:81–90, 2015.
- Gina R. Kuperberg and T. Florian Jaeger. What do we mean by prediction in language comprehension? *Language, cognition and neuroscience*, 31(1):32–59, 2016.
- Hope Landrine. Race \times class stereotypes of women. *Sex roles*, 13(1-2):65–75, 1985.
- Gene H. Lerner. Finding “face” in the preference structures of talk-in-interaction. *Social psychology quarterly*, pages 303–321, 1996.
- Stephen C. Levinson. Speech acts. In *Oxford handbook of pragmatics*, pages 199–216. Oxford University Press, 2017.
- Lilla Magyari and Jan P. de Ruiter. Prediction of turn-ends based on anticipation of upcoming words. *Frontiers in psychology*, 3:376, 2012.
- Lilla Magyari, Jan P. de Ruiter, and Stephen C. Levinson. Temporal preparation for speaking in question-answer sequences. *Frontiers in Psychology*, 8:211, 2017.
- Nobuko Osada. Listening comprehension research: A brief review of the past thirty years. *Dialogue*, 3(1):53–66, 2004.
- Tepring Piquado, Derek Isaacowitz, and Arthur Wingfield. Pupillometry as a measure of cognitive effort in younger and older adults. *Psychophysiology*, 47(3):560–569, 2010.
- Anita Pomerantz. Agreeing and disagreeing with assessments: Some features of preferred/dispreferred turn shaped. In *Structures of social action: Studies in conversation analysis*. Cambridge University Press, 1984.
- Seán G. Roberts, Francisco Torreira, and Stephen C. Levinson. The effects of processing and sequence organization on the timing of turn taking: a corpus study. *Frontiers in psychology*, 6: 509, 2015.
- Jeffrey D. Robinson. *Accountability in social interaction*. Oxford University Press, 2016.
- Jeffrey D. Robinson and Heidi Kevoe-Feldman. Using full repeats to initiate repair on others’ questions. *Research on Language and Social Interaction*, 43(3):232–259, 2010.
- Christoph Rühlemann. How long does it take to say ‘well’? evidence from the audio bnc. *Corpus Pragmatics*, 3(1):49–66, 2019.

- Harvey Sacks, Emanuel A. Schegloff, and Gail Jefferson. A simplest systematics for the organization of turn taking for conversation. In *Studies in the organization of conversational interaction*, pages 7–55. Elsevier, 1978.
- Emanuel A. Schegloff. Sequencing in conversational openings 1. *American anthropologist*, 70(6): 1075–1095, 1968.
- Emanuel A. Schegloff. The relevance of repair to syntax-for-conversation. In *Discourse and syntax*, pages 261–286. Brill, 1979.
- Emanuel A. Schegloff. Discourse as an interactional achievement: Some uses of ‘uh huh’ and other things that come between sentences. *Analyzing discourse: Text and talk*, 71:93, 1982.
- Emanuel A. Schegloff. *Sequence organization in interaction: A primer in conversation analysis I*, volume 1. Cambridge university press, 2007.
- Emanuel A. Schegloff, Gail Jefferson, and Harvey Sacks. The preference for self-correction in the organization of repair in conversation. *Language*, 53(2):361–382, 1977.
- Bruce A. Schneider, Liang Li, and Meredyth Daneman. How competing speech interferes with speech comprehension in everyday listening situations. *Journal of the American Academy of Audiology*, 18(7):559–572, 2007.
- Saul Sternberg, Stephen Monsell, Ronald L. Knoll, and Charles E. Wright. The latency and duration of rapid movement sequences: Comparisons of speech and typewriting. In *Information processing in motor control and learning*, pages 117–152. Elsevier, 1978.
- Tanya Stivers and Jeffrey D. Robinson. A preference for progressivity in interaction. *Language in society*, pages 367–392, 2006.
- Tanya Stivers, Nicholas J. Enfield, Penelope Brown, Christina Englert, Makoto Hayashi, Trine Heinemann, Gertie Hoymann, Federico Rossano, Jan P. de Ruiter, Kyung-Eun Yoon, et al. Universals and cultural variation in turn-taking in conversation. *Proceedings of the National Academy of Sciences*, 106(26):10587–10592, 2009.
- JASP Team. JASP (version 0.10.0). *Computer software*, 2018. URL <https://jasp-stats.org>.
- Walter J. B. van Heuven, Pawel Mandera, Emmanuel Keuleers, and Marc Brysbaert. SUBTLEX-UK: A new and improved word frequency database for british english. *Quarterly journal of experimental psychology*, 67(6):1176–1190, 2014.
- Jeffrey J. Walczyk, Karen S. Roper, Eric Seemann, and Angela M. Humphrey. Cognitive mechanisms underlying lying to questions: Response time as a cue to deception. *Applied Cognitive Psychology*, 17(7):755–774, 2003. doi: 10.1002/acp.914.
- Candace West and Don H. Zimmerman. *Small insults: A study of interruptions in cross-sex conversations between unacquainted persons*. Critical concepts in psychology. Routledge/Taylor & Francis Group, New York, NY, US, 2015. ISBN 978-0-415-66094-5 (Hardcover).
- Victor H. Yngve. On getting a word in edgewise. In *Chicago Linguistics Society, 6th Meeting, 1970*, pages 567–578, 1970.