2021

Examination of Conceptual Pacts and Perspective Taking in Remote Communication

Ashley Pieper

FLORIDA STATE UNIVERSITY

COLLEGE OF ARTS AND SCIENCES

EXAMINATION OF CONCEPTUAL PACTS AND PERSPECTIVE TAKING IN REMOTE COMMUNICATION

By

ASHLEY PIEPER

A Thesis submitted to the
Department of Psychology
in partial fulfillment of the
requirements for the degree of
Master of Science

2021

Ashley Pieper defended this thesis on October 1, 2021.
The members of the supervisory committee were:

Michael P. Kaschak
Professor Directing Thesis

Sara Hart
Committee Member

Walter R. Boot
Committee Member

The Graduate School has verified and approved the above-named committee members, and certifies that the thesis has been approved in accordance with university requirements.

ACKNOWLEDGMENTS

I would like to thank my advisor, Dr. Michael Kaschak for helping guide me through the research process and always having his door open for my many questions. His knowledge and kindness have been essential to my academic journey. I would also like to thank my committee members, Dr. Wally Boot and Dr. Sara Hart for taking the time to read this thesis and giving their feedback and support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Conceptual pacts are agreed-upon ways of talking about things that are specific to a conversation and a conversation partner. Most research focuses on conceptual pacts with participants being co-present in some way. However, we do not know if previous findings still apply when participants are not co-present. Along similar lines, Krauss and colleagues found that participants do still take their partner into account in asynchronous studies by examining the differences in written letters made by participants with different people in mind – themselves, a friend, and a stranger (Fussell & Krauss, 1989; Krauss, 1987; Krauss, Vivekananathan, & Weinheimer, 1968). However, these study did not look at conceptual pacts. The present series of studies aimed to deepen the areas understanding of conceptual pacts and asynchronous conversations by expanding on Krauss' work to include conceptual pacts in asynchronous settings. Overall, our data did not support that conceptual pacts are made during asynchronous conversations. These findings suggest follow-up studies to clarify in what settings conceptual pacts are made.

CHAPTER 1

INTRODUCTION

At its core, language is a social activity. This is true for children, who learn language by interacting with their parents, siblings, and other caregivers (e.g., Tomasello, 2003), and for adults, whose language use largely consists of interacting with the people around them (e.g., Dunbar, 1998). In order for language to facilitate communication effectively, there needs to be common ground between people (e.g., Clark, 1996). Common ground can be thought of as shared experiences, points of view, attitudes, and knowledge between two or more people. During a conversation, the extent to which common ground is (or is not) present can affect a speaker's behavior in many ways: for example, deciding what to say, how to say it, and determining how to interpret their partner's utterances (e.g., Clark, 1996).

Although few researchers doubt that common ground is a crucial aspect of conversation and social interaction in general, there are different ideas about when a speaker would use common ground during a conversation. Are people generally more egocentric or allocentric when it comes to communication? Some researchers believe that people are generally egocentric and only take their conversation partner's perspective into account when necessary (Pickering and Garrod, 2004; Keysar, Barr, and Horton, 1998). In this view, a person does not actively try to represent the contents of their partner's mind, and common ground is limited to what the speaker knows (e.g., general assumptions about language users or specific memories of their conversation partner). The only exceptions to this are cases where something breaks down in the conversation (e.g., there is a misunderstanding), and the speaker tries to take the listener's perspective into account to resolve the issue. The alternative view is that people always consider their partner's perspective during a conversation and attempt to tailor it accordingly (e.g.,

Metzing and Brennan, 2003). The common ground thus not only includes the individual speaker's knowledge about the current context and their partner, but also an active consideration of what is in their partner's mind.

Common ground can occur in a conversation in different ways. It may be brought into the conversation or built during it. For example, a conversation with a stranger may consist of general knowledge about the world (e.g., the weather, recent events in the area, the red car across the street), but with a friend could be more personalized (e.g., an event that occurred during a dinner together, a mutual friend). However, common ground can also be built during a conversation. Those two strangers mentioned earlier may have started talking about the weather and eventually discover they both are from the same state. They may also develop a new, more specific code to refer to particular elements of their world (e.g., deciding that they will call that car parked across the street a mustang or a red car). This kind of ad hoc communicative innovation is called a conceptual pact (Brennan and Clark, 1996). Conceptual pacts are common ground that we make up on the fly during a conversation. They are agreed-upon ways of talking about things that are specific to a conversation and a conversation partner. Conceptual pacts help facilitate interactions and make them more efficient by creating common ground between conversation partners.

Several researchers have documented the development of conceptual pacts between speakers (e.g., Brennan and Clark, 1996; Branigan, Pickering, and Cleland, 2000; Metzing and Brennan, 2003). These studies virtually always involve conversation partners that are co-present in the lab. Under these conditions, researchers have shown that conversation partners readily establish conceptual pacts (Brennan and Clark, 1996) and that these pacts are partner-specific (e.g., Metzing & Brennan, 2003). If A and B have been engaged in the task and establish

conceptual pacts, and then A starts interacting with C, A will not assume that C shares the same knowledge as B or that the conceptual pacts with B still hold. However, if B is reintroduced into the dyad, A and B will readily resume use of their conceptual pacts. There has been considerably less research done when participants are not co-present and are not communicating in real-time. In a series of studies from Robert Krauss' lab (Fussell & Krauss, 1989; Krauss, 1987; Krauss, Vivekananathan, & Weinheimer, 1968), participants were asked to imagine interacting with another person (themselves, a friend, and a stranger) while writing a letter. The content of the letters was different based on the intended reader, suggesting that people can take non-present partners into account when generating their responses and supported that humans are allocentric when it comes to common ground in communication (Fussell & Krauss, 1989; Krauss, 1987; Krauss, Vivekananathan, & Weinheimer, 1968). However, this work did not specifically examine conceptual pacts (the participant only made one communication to the non-present partner, so did not need to establish a conceptual pact).

The present study extends Krauss' work to examine whether individuals can generate conceptual pacts with a non-present partner. Two experiments were conductedIn both experiments participants were presented with ten hard-to-describe shapes, called tangrams. Tangrams have been used in similar studies in past research (Brennan & Clark, 1996; Clark & Wilkes-Gibbs, 1986; Schober & Clark, 1989; Wilkes-Gibbs & Clark, 1992). Participants were asked to describe these so that they could be given to a partner who would use the descriptions to select the correct item out of the set of tangrams. The participants were either told they would have the same partner across multiple rounds of the game or that they would have a different partner for each round of the game If the participant is modeling the mind of their non-present partner, following the allocentric view, we expect that when participants have the same partner

repeatedly, they will show a) more consistent descriptions across rounds of the game (since it would be easiest to keep the descriptions the same if you have the same partner), and b) progressively shorter descriptions across trials (because the establishment of conceptual pacts makes it easier to have short-hand descriptions of the tangrams). The egocentric view, in contrast, would hypothesize that none of these would happen.

To test these hypotheses, we measured Word Count and two types of changes. The first type of changes measured were Strict Changes. These changes consisted of any change in the participants' description from one run to the next. For example, making a small change from "man" to "person" or a large change such as "cliff" to "diamond". The second type of changes measured were Loose Changes. These changes were only large changes in the descriptions. For example, changing from "cliff" to "diamond" would count as a Strict Change and a Loose Change. However, changing from "man" to "person" would not count as a Loose Change.

CHAPTER 2

EXPERIMENT 1

Method

Participants

The participants were 61 undergraduate psychology students from Florida State University. They participated in exchange for extra credit or course credit. Before coding and analyzing the data, we eliminated participants who skipped more than 3 of the ten tangrams in each of the three rounds of the task (see below for details). We excluded 16 participants on this basis, leaving a final sample of 45 participants. This sample size is similar to past research.

Materials

Ten tangram shapes were used in each of the three runs (see image 1).

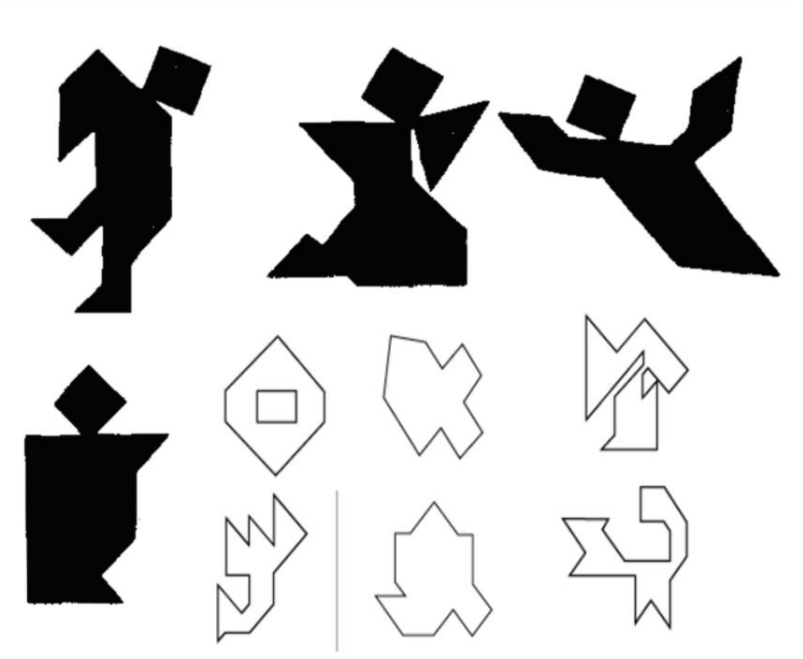

Figure 1. Ten tangrams used in the study.

Procedure

Florida State University's Institutional Review Board approved this project. The experiment was conducted online via qualtrics in June and July of 2020.

Participants were randomly assigned to one of 2 conditions. In the first condition, participants were told they would have the same partner for all runs in the study. In the second condition, they were told they would have a different partner for each run.

Both conditions consisted of three runs. Each run showed all ten tangrams (see materials) in random order. Participants were asked to describe each tangram so that their partner, the matcher, would be able to pick the tangram out of a set. Between each run, in both conditions, participants were reminded of who their next partner would be.w

On each page the participant saw one tangram with a typing box below. To move on to the next tangram, they had to click "next" in the bottom corner of their screen.

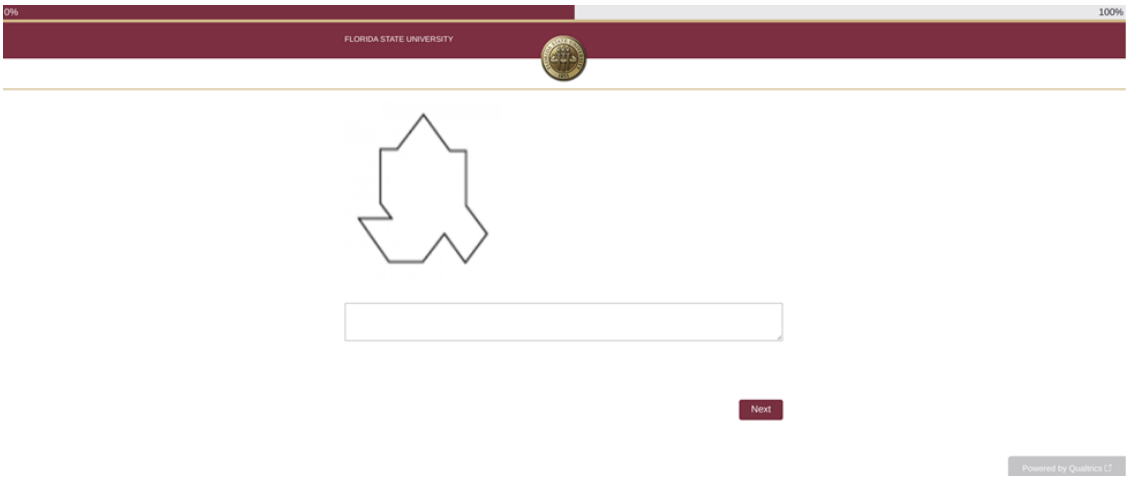

Figure 2. Example of participant screen.

The instructions for both the “same partner” condition and the “different partner” condition are available on the Open Science Framework (same partner: https://osf.io/psq8m/; different partner: https://osf.io/85mb2/).

Coding

Word Count. Excel was used to generate the Word Count for every response. The following code was used to generate the counts:

=IF(ISBLANK(responseblock),0,LEN(TRIM(responseblock))-
LEN(SUBSTITUTE(TRIM(Tresponseblock)," ",""))+1)

Strict Changes. For this study, “Strict Changes” refers to any change in the participants’ description for a tangram from one run to the next. For example, if a participant called a tangram a “man with a hat on” in run 1, and changed to “a person with a hat on” in run 2, that would count as a Strict Change even though a “man” and a “person” have essentially the same meaning. Furthermore, if a participant called something a “man with a hat on” in the first run, and changed to a “cliff with a box” in the second, that was also coded as a Strict Change because it is also not the same wording in both.

Loose Changes. Loose Changes for this study were any changes in the description from one run to another that go from one general topic to something entirely different. Going back to the examples above, if a participant called a tangram a “man with a hat on” in run 1, and changed to a “a person with a hat on” in run 2, that was not counted as a Loose Change because a “man” and a “person” have essentially the same meaning. However, if a participant called something a

"man with a hat on" in the first run, and changed to a "cliff with a box" in the second, that was coded as a Loose Change because those two descriptions are not the same general idea.

Reliability. Three DIS students in the lab assisted in coding for reliability. Each coded 20% of the participant responses and compared them to the main researcher's coding. The first coding was found to be unreliable. The averages of DIS 1 and the researcher was .699 (Strict = .330, Loose = .737). Due to the large discrepancy between the coding reliability in each condition, all participant responses were re-coded by the main researcher and two new DIS students were recruited. This run of coding was found to be reliable and was used to analyze the data. The average reliability of all DIS and the researchers coding was .785 (DIS 1: Strict = coder had no variability, Loose = .806; DIS 2: Strict: .845, Loose = .704). Due to the subjective nature of the coding, this reliability rating was acceptable. We believe the reliability was higher due to new DIS students paying closer attention to the coding and the main researcher having more experience after going through and coding once.

Design and Analysis

Word Count, Strict Changes, and Loose Changes were analyzed using mixed models regression. The model for Word Count used Condition (1 = different partner, -1 = same partner), Time (cycles through the task coded as 1, 2, or 3), and the interaction of these variables as predictors. The model included Participant ID and Items as crossed random factors and random slopes for Time across Participants and Items and Condition across Items.

The model for the Strict Changes used Condition (1 = different partner, 0 = same partner) as the predictor. The model included Participants and Items as crossed random factors and the

random slope for Condition across Items. The model for Loose Changes had the same structure as the model for the Strict Changes.

All analyses were conducted using the glmertree package (Fokkema M, Smits N, Zeileis A, Hothorn T, Kelderman H, 2018) in RStudio (R Core Team, 2013). The data and code used for these analyses are available on the Open Science Framework (Word Count: https://osf.io/62t7a/; Strict Coding and Loose Coding: https://osf.io/h45vt/).

Results and Discussion

Strict Changes

Table 1. Results, means, and standard deviations for Strict Changes.

Predictor	Estimate	SE	*p*
Intercept	0.945	0.125	<.001
Condition	-0.023	0.151	0.876
Means and Standard Deviation of Conditions			
	Mean	Standard Deviation	
Same Partner	0.527	0.711	
Different Partner	0.639	0.709	

The regression analysis and descriptive statistics for the Strict Changes are shown in Table 1. Participants in the Different Partner condition produced more Strict Changes than participants in the Same Partner condition. However, this difference was not statistically reliable (p = .876).

Loose Changes

Table 2. Results, means, and standard deviations for Loose Changes.

Predictor	Estimate	SE	*p*
Intercept	0.382	0.112	0.002
Condition	0.183	0.144	0.210

Table 2. - continued

Predictor	Estimate	SE	*p*
Means and Standard Deviation of Conditions			
	Mean	Standard Deviation	
Same Partner	0.241	0.516	
Different Partner	0.426	0.599	

Table 2 shows the regression analysis and descriptive statistics for the Loose Changes. Participants in the Different Partner condition produced more Loose Changes than participants in the Same Partner condition, but this difference was not statistically reliable ($p = .210$).

Word Count

Table 3. Results, means, and standard deviations for Word Count.

Predictor	Estimate	SE	*p*
Intercept	15.204	2.429	<.001
Condition	-4.479	2.399	0.069
Time	-1.449	0.456	0.003
Condition:Time	0.699	0.456	0.132
Means and Standard Deviation of Conditions			
	Mean	Standard Deviation	
Time 1			
Same Partner	17.386	19.358	
Different Partner	9.930	8.909	
Time 2			
Same Partner	15.672	19.561	
Different Partner	9.309	9.238	
Time 3			
Same Partner	13.091	16.405	
Different Partner	8.430	9.555	

The regression analysis and descriptive statistics for Word Count are shown in Table 3. The only significant effect was the main effect of Time ($p = .003$): participants produced fewer words as they moved through the runs of the experiment. Across runs of the experiment,

participants in the Same Partner condition shortened their descriptions more than participants in the Different Partner condition. However, the Time x Condition interaction was not significant (p = .132).

Although the means presented in Tables 1, 2, and 3 generally follow the predicted patterns, none of the critical effects were statistically reliable. The weakness of our results may have been due to a couple of factors. First, it is possible that some of the participants did not thoroughly read the instructions and may not have paid attention to whether they were interacting with the same or different partners. In addition, the sample size of the study was relatively small (n=23 in each group). In Experiment 2, we fixed these issues by making the participants answer whether they would have the same or different partner before beginning any of the runs and **bolding** the part of the instructions that specified this information. Images of their "partners" added to the instruction page further ensured they paid attention to who their partner was. Lastly, we employed Frick's (1998) sequential sampling method to ensure we had enough participants.

CHAPTER 3

EXPERIMENT 2

Introduction

Experiment 1 found general (but non-significant) trends toward more Strict Changes and Loose Changes in the Different Partner condition compared to the Same Partner condition. In addition, Word Count tended to decrease more sharply in the Same Partner condition than in the Different Partner condition. However, the structure of Experiment 1 caused concern about whether participants paid enough attention to the nature of their partner and about the sample size. In Experiment 2, we fixed those issues by adding an image of the "partner" to the instruction page, by making the participants choose either a "same partner" or "different partner" answer choice before beginning any of the runs, and by bolding the part of the instructions that specified this information.

To fix the issue of the small sample size, we employed Frick's (1998) sequential sampling method. In this method, the data from the experiment are analyzed after collecting responses from a predetermined number of participants. If the p-values are > .36, data collection stops, and the effects are considered non-significant. If the p-values are < .01, data collection stops, and the effects are considered significant. If the p-values are between .36 and .01, another batch of participants is recruited. The process continues until the results become conclusive ($p > .36$ or $p < .01$), or the number of participants reaches a predetermined maximum. Frick's (1998) method allows for sequential sampling while keeping the family-wise alpha level at .05.

To further verify any changes made were due to conceptual pacts being made, we also added a fourth run. In this run, participants in both conditions were told they would have a new partner. We expected that there would be no change to the Different Partner condition, as they

had a new partner in every run. However, we expected that there would be an increase in Strict Changes, Loose Changes, and Word Count in the Same Partner condition. Researchers have found that conceptual pacts made with one person do not transfer to another - participants do not expect a new partner to know their conceptual pacts with their old partner (Brennan & Clark, 1996; Wilkes-Gibbs & Clark, 1992; Schober & Clark 1989). Therefore, if the participant created a conceptual pact with their original partner, they would not apply it to their new partner.

Method

Participants

We used Frick's (1998) sequential sampling method to reach our final sample size for Experiment 2. We collected an initial sample of 143 undergraduate students from Florida State University. Participants were excluded if they met any of the following criteria: did not answer 7 out 10 questions for each run, stated they typed less because they were tired of typing, or correctly guessed the aim of the study. 31 participants were excluded from this collection due to meeting one of these criteria, leaving a total of 112 participants. Data analysis revealed p-values between .36 and .01. Therefore, we collected another round of participants. We aimed for 60 participants (30 in each condition) and gathered 74. Due to the criteria mentioned previously, 34 participants were excluded, leaving 40 participants. Our p-values were again between .36 and .01. Therefore, we collected a final sample of 57 participants. 28 were excluded, leaving us with 29 participants. This gave us a total of 181 participants (where our original intended maximum was 180). Therefore, data collection stopped at this point. Participants received extra credit or course credit for participating.

Materials

The materials used were the same as in Experiment 1 in addition to images of the "participants" that were added to the instruction pages.

Procedure

The experiment was conducted online via Qualtrics in January, February, and July of 2021.

The procedure used was the same as Experiment 1, with some minor changes. The first change was that there were four trials total rather than three. For the last trial, both the Same Partner condition and the Different Partner condition were told they would have a different partner. Furthermore, after reading the beginning instructions, as well as between each run, participants were asked to acknowledge whether they would have the same partner or a different partner by clicking the correct answer choice before moving on to the next run. The rest of the procedure was the same as in experiment 1.

Coding

The data was coded as in Experiment 1 with the addition of coding both Strict Changes and Loose Changes from runs 3 to 4.

Reliability. Three DIS students were recruited to code 20% of the participant responses to test for reliability. These DIS' coding was compared to the main researchers coding, similar to Experiment 1. Due to using Frick's (1998) sequential sampling method, we had multiple runs of participants. For the first run of participant data the average reliability was .670 (Coder 1: Strict 1-3 = coder had no variability, Loose 1-3 = .423, Strict 3-4 = coder had no variability, Loose 3-4

= .633; Coder 2: Strict 1-3: .894, Loose 1-3 = .888, Strict 3-4 = .643, Loose 3-4 = .774; Coder 3: Strict 1-3 = .721, Loose 1-3 = .699, Strict 3-4 = .694, Loose 3-4 = .53). Due to the subjective nature of the coding, this reliability rating was found to be acceptable.

In the second run of coding, two of the coders from run 1 returned. However, the coders seemed to produce many subjective errors and did not apply the instruction standards in this run (A participant stating "a baseball field" for all four times seeing the same image, and a coder coding two changes). The average reliability was .562 (Coder 1: Strict 1-3 = coder had no variability, Loose 1-3 = .62, Strict 3-4 = coder had no variability, Loose 3-4 = .683; Coder 3: Strict 1-3: .857, Loose 1-3 = .359, Strict 3-4 = .74, Loose 3-4 = .101). Therefore, another DIS coder was recruited. The average reliability for this coder was .763 (Coder 4: Strict 1-3: .853, Loose 1-3 = .752, Strict 3-4 = .832, Loose 3-4 = .614). The addition of this coder brought the overall average reliability for the second run up to .639. Due to the subjective nature of the coding, this reliability rating was acceptable.

Design and Analysis

Data analysis was similar to Experiment 1 for runs 1 through 3. However, another analysis was added comparing runs 3 and 4. This analysis was also similar to Experiment 1 in setup. This additional analysis used mixed models regression, similar to Experiment 1. The data and code used for these analyses are available on the Open Science Framework (Word Count: https://osf.io/g7var/; Strict and Loose Coding: https://osf.io/rjnxs/).

Results and Discussion

Strict Changes

Table 4. Results, means, and standard deviations for Strict Changes runs 1 through 3.

Predictor	Estimate	SE	*p*
Intercept	1.044	0.060	<.001
Condition	-0.028	0.045	0.536
Means and Standard Deviation of Conditions			
	Mean	Standard Deviation	
Same Partner	1.017	0.861	
Different Partner	1.068	0.853	

Table 4 lists the results of the analysis of Strict Changes for runs 1 through 3. Similar to Experiment 1, participants in the Different Partner conditions were expected to have more Strict Changes in their responses than those in the Same Partner condition from runs 1 through 3. This prediction was not supported. The effect of Condition on Strict Changes for runs 1 through 3 was not significant (*p*=0.536).

Table 5. Results, means, and standard deviations for Strict Changes runs 3 to 4.

Predictor	Estimate	SE	*p*
Intercept	0.406	0.030	<.001
Condition	-0.026	0.026	0.319
Means and Standard Deviation of Conditions			
	Mean	Standard Deviation	
Same Partner	0.380	0.486	
Different Partner	0.430	0.495	

Strict Changes from run 3 to 4 were analyzed on the expectation that there would be an increase in changes in the Same Partner condition, as they went from three runs with the same partner to a new partner in the 4th run. However, this was not supported. Table 4 lists the results

of the analysis of Strict Changes for runs 3 through 4. The effect of Condition on Strict Changes for runs 3 through 4 was not significant *(p*=.319).

Loose Changes

Table 6. Results, means, and standard deviations for Loose Changes runs 1 through 3.

Predictor	Estimate	SE	*p*
Intercept	0.526	0.048	<.001
Condition	-0.020	0.036	0.573
Means and Standard Deviation of Conditions			
	Mean	Standard Deviation	
Same Partner	0.506	0.741	
Different Partner	0.545	0.799	

Table 6 lists the results of the analysis of Loose Changes for runs 1 through 3. Similar to Experiment 1, participants in the Different Partner condition were expected to have more changes in their responses than those in the Same Partner condition. This was not supported. The effect of Condition on Loose Changes for runs 1 through 3 was not significant *(p*=.573).

Table 7. Results, means, and standard deviations for Loose Changes runs 3 to 4.

Predictor	Estimate	SE	*p*
Intercept	0.190	0.021	<.001
Condition	-0.026	0.019	0.174
Means and Standard Deviation of Conditions			
	Mean	Standard Deviation	
Same Partner	0.216	0.370	
Different Partner	0.164	0.511	

Table 7 lists the results of the analysis of Loose Changes for runs 3 through 4. Loose Changes from run 3 to 4 were analyzed on the expectation that there would be an increase in

changes in the Same Partner condition. However, this was not supported. The effect of Condition on Loose Changes for runs 3 through 4 was not significant *(p*=.174).

Word Count

Table 8. Results, means, and standard deviations for Word Count for trials 1 through 3.

Predictor	Estimate	SE	*p*
Intercept	20.920	1.316	<.001
Condition	-0.849	1.180	0.473
Time	-2.300	0.170	<.001
Condition:Time	0.553	0.170	0.001
Means and Standard Deviation of Conditions			
	Mean	Standard Deviation	
Time 1			
Same Partner	18.514	19.519	
Different Partner	19.685	23.968	
Time 2			
Same Partner	16.198	17.062	
Different Partner	14.962	15.089	
Time 3			
Same Partner	15.020	15.562	
Different Partner	13.981	16.730	

Word count was expected to decrease significantly more in the Same Partner condition than in the Different Partner condition. Furthermore, it was also predicted to decrease for both conditions across trials. The results of this analysis are listed in Table 8. The effect of Condition was not significant (p=0.473). However, the effect of Time ($p < .001$) and the interaction of Time and Condition (p=0.001) were significant. Participants generally produced fewer words as they progressed through the experiment, and contrary to predictions, this tendency was more pronounced for participants in the Different Partner condition than participants in the Same Partner condition.

Table 9. Results, means, and standard deviations for Word Count for trials 3 to 4.

Predictor	Estimate	SE	*p*
Intercept	16.541	1.395	<.001
Condition	3.313	1.325	0.013
Time	-0.702	0.245	0.004
Condition:Time	-0.910	0.245	0.0002
Means and Standard Deviation of Conditions			
	Mean	Standard Deviation	
Time 3			
Same Partner	15.020	15.562	
Different Partner	13.981	16.730	
Time 4			
Same Partner	11.384	13.234	
Different Partner	14.789	17.400	

Word Count from run 3 to 4 was analyzed on the expectation that there would be an increase in word count in the Same Partner condition, but the Different Partner condition would not change. Table 9 lists these results. While we did find a significant interaction between Time and Condition (p=0.0002), it was the opposite of what we expected. Same Partner participants produced fewer words in time 4 than in time 3, and Different Partner participants produced more words in time 4 than in time 3. Furthermore, the effects of both Condition (p = .013) and Time (p = .004) were significant.

CHAPTER 3

GENERAL DISCUSSION

These experiments aimed to expand on Robert Krauss's findings on partner-specific language, as well as the extensive research on conceptual pacts, to include asynchronous interactions (e.g., Fussell & Krauss, 1989; Krauss, 1987; Krauss, Vivekananathan, & Weinheimer, 1968; Brennan and Clark, 1996; Metzing & Brennan, 2003). Robert Krauss and associates examined the writing style of participants in reference to who would later be reading it - themselves, a stranger, or a friend. They found that participants' writing differed significantly for each group, suggesting that humans can take non-present partners into account when interacting asynchronously. Furthermore, researchers have found that, similar to the studies done by Krauss, conceptual pacts are partner-specific (e.g., Metzing & Brennan, 2003). Combining the findings of Krauss' research with the extensive research on conceptual pacts led us to conduct the present study. We looked to see if conceptual pacts could be made by participants in an asynchronous setting.

If our participants were behaving as if they had made a conceptual pact with their partner (in the Same Partner condition), we expected that a) participants in the Same Partner condition would shorten their descriptions across runs of the experiment more than participants in the Different Partner condition, b) participants in the Same Partner condition would make fewer Loose Changes and Strict Changes to their descriptions, and c) in Experiment 2, in the transition from the 3rd to the 4th run of the experiment, participants in the Same Partner condition who were told that they now had a different partner would show an increase in the number of changes they made to their descriptions and an increase in the number of words in their descriptions. These predictions were not supported.

Our results align with the view that language users are egocentric (e.g., Keysar et al., 1998). In other words, language users do not consider the needs of their conversation partner (or audience) unless there is a specific need to do so. Brennan and Clark (1996) and Krauss and colleagues (e.g., Fussell and Krauss, 1989) demonstrate that language users consider their audience both in one-shot communications (Krauss and colleagues' work) and in dialogues (Brennan & Clark, 1996). It is important to note that, in these cases, the design of the experiment was such that the participants had information about their audience that could change their language choices. For example, in Krauss and colleagues' work, the participant was either very familiar with the target of their writing (themselves, or a friend) or not, and could use this information to make decisions about what to write (e.g., they knew enough about their friend to tailor their message appropriately). In Brennan and Clark's (1996) study, participants interacted with a partner. Therefore, they would know if their description of a target object was not successful and could negotiate a better label with their partner. In our studies, participants did not know anything about their partners and did not have the opportunity to get feedback about their descriptions. This absence of feedback may explain why the participants who were told they would have the same partner throughout the experiment did not behave all that differently from the participants who were told they would have a different partner. Relative to those with a variety of partners, being aware that they had the same partner through the experiment was not enough information to cause a change to the descriptions in systematic ways.

Our participants' lack of knowledge about their partners and their partners responses to their descriptions is likely an important factor in driving similar behavior in the Same and Different Partner conditions. Under these circumstances, a reasonable strategy for both groups of participants to use would be to find a description that they thought was reasonable and then to

continue using that description regardless of their partner. Thus, where we expected that participants in the Same Partner condition would be more stable in their descriptions because they wanted to establish consistency with their partner, it may be the case that being consistent across trials was also easier for the participants in the Different Partner condition. It remains to be seen whether providing our participants with more information about their partners (e.g., knowledge about their hobbies that could be used to shape the tangram descriptions) would create a situation where the participants tailored their descriptions to the partner enough (i.e., established a conceptual pact) that they would need to change their descriptions when writing for a different partner.

To conclude, we did not find evidence that giving our participants knowledge about whether they were interacting with the same partner or a different partner affected the pattern of their tangram descriptions. Future explorations of this topic might manipulate the knowledge that participants have about their partners further, examining under what circumstances asynchronous conversation behaves similarly to synchronous interactions.

APPENDIX A

IRB EXEMPTION

FLORIDA STATE UNIVERSITY
OFFICE *of the* VICE PRESIDENT *for* RESEARCH

EXEMPTION DETERMINATION

January 29, 2020

Ashley Pieper,

Dear Ashley Pieper:

On 1/29/2020, the IRB staff reviewed the following submission:

Type of Review:	Exempt (3)(i)(B) Benign behavioral interventions (low risk)
Title:	Social Factors of Dialogue
Investigator:	Ashley Pieper
Submission ID:	STUDY00000930
Study ID:	STUDY00000930
Funding:	None
Grant ID:	None
IND, IDE, or HDE:	None
Documents Reviewed:	• Social Factors of Dialogue Debriefing.pdf, Category: Debriefing; • Social Factors of Dialogue Debriefing Deception.pdf, Category: Debriefing; • Social Factors of Dialogue Information Sheet Deception.pdf, Category: Consent Form; • Social Factors of Dialogue Information Sheet.pdf, Category: Consent Form;

The IRB staff determined the protocol qualifies for exemption, effective on 1/29/2020.

You are advised that any modification(s) to the protocol for this project that may alter this exemption determination must be reviewed and approved prior to implementation of the proposed modification(s).

Modifications to the research may invalidate the exemption determination (because the research no longer meets the exemption criteria described in HRP-312 – WORKSHEET – Exemption Determination).

Examples of minor changes to exempt research that would *not* alter the exemption determination and should therefore not be submitted to the IRB for further review include the following:

- Making administrative (formatting, grammar, spelling) revisions to the protocol, consent or recruitment materials or other study documents
- Adding or revising non-sensitive questions or non-identifiable response options to a survey, interview, focus group or other data collection instrument

- Increasing or decreasing the number of study subjects—*unless* adding a new study sample such as children or prisoners or adding a new source of data or records
- Making study team/personnel changes—*except* a change in Principal Investigator (PI)

Examples of changes to exempt research that ***do require*** prospectively submitting a modification to the IRB before implementing changes include the following:

- Making substantive revisions or additions (e.g., change in PI; funding source; sample; source of study subjects or their data; study sites or settings; procedures, interventions or interactions with study subjects; use of any drug, device, supplement or biologic; study subjects' time or duration spent performing or participating in study activities) to the protocol, consent or recruitment materials or other study documents
- Adding or revising sensitive questions or identifiable response options to a survey, interview, focus group or other data collection instrument
- Adding a new study sample such as children or prisoners or adding a new source of data or records
- Obtaining, using, studying, analyzing, generating, storing or maintaining identifiable information or identifiable biospecimens in addition to or in lieu of de-identified or anonymous information or specimens
- Change in study risks (e.g., impact upon study subjects; impact upon students' opportunity to learn educational content or assessment of educators who provide instruction; any disclosure of study subjects' responses outside of the research may place study subjects at risk of criminal or civil liability or be damaging to subjects' financial standing, employability, educational advancement or reputation)
- Change in Principal Investigator (PI) or (for students) faculty advisor
- New or change in financial interest

In conducting this protocol, you are required to follow the applicable requirements listed in the Investigator Manual (HRP-103), which can be found by navigating to the Library within the RAMP IRB system.

Sincerely,

Office for Human Subjects Protection (OHSP)
Florida State University Office of Research
2010 Levy Avenue, Building B Suite 276
Tallahassee, FL 32306-2742
Phone: 850-644-7900
OHSP Group Email: humansubjects@fsu.edu
OHSP Web: https://www.research.fsu.edu/hs

APPENDIX B

INSTRUCTION SHEET

FLORIDA STATE UNIVERSITY

Study Title: Social Factors of Dialogue

Principal Investigator: Ashley Pieper

You are being asked to voluntarily participate in a research study. We are doing this study in an effort to learn more about dialogue and conversation. If you choose to participate, you will be asked to describe a series of images as you see them.

Your participation in this research is completely voluntary, and you may stop the study at any time without penalty. If you are participating in this research study as part of a class, you can ask your instructors about alternative activities that you may use in place of research participation.

Your decision whether or not to participate will not affect your current or future relations with the University. If you decide to participate, you are free to withdraw at any time without affecting those relationships.

If you have any questions, please contact Ashley Pieper

If you have any questions or concerns about your rights as a research participant, or regarding the study and would like to talk to someone other than the researcher(s), you are encouraged to contact the FSU IRB at telephone number 850-644-7900. You may also contact this office by email at humansubjects@fsu.edu, or by writing or in person at 2010 Levy Street, Research Building B, Suite 276, FSU Human Subjects Committee, Tallahassee, FL 32306-2742.

REFERENCES

Branigan, H. P., Pickering, M. J., & Cleland, A. A. (2000). Syntactic co-ordination in dialogue. Cognition, 75(2), B13-B25.

Brennan, S. E., & Clark, H. H. (1996). Conceptual pacts and lexical choice in conversation. Journal of Experimental Psychology: Learning, Memory, and Cognition, 22(6), 1482.

Clark, H. H. (1996). Using language. Cambridge university press.

Clark, H. H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. Cognition, 22(1), xy1-39.

Dunbar, R., & Dunbar, R. I. M. (1998). Grooming, gossip, and the evolution of language. Harvard University Press.

Fokkema M, Smits N, Zeileis A, Hothorn T, Kelderman H (2018). "Detecting Treatment-Subgroup Interactions in clustered Data With Generalized Linear Mixed-Effects Model Trees." Behavior Research Methods, 50, 2016-2034. http://link.springer.com/article/10.3758/s13428-017-0971-x.

Frick, R. W. (1998). A better stopping rule for conventional statistical tests. Behavior Research Methods, Instruments, & Computers, 30, 690-697.

Fussell, S. R., & Krauss, R. M. (1989). Understanding friends and strangers: The effects of audience design on message comprehension. European Journal of Social Psychology, 19(6), 509-525.

Keysar, B., Barr, D. J., & Horton, W. S. (1998). The egocentric basis of language use: Insights from a processing approach. Current Directions in Psychological Science, 7, 46–50.

Krauss, R. M. (1987). The role of the listener: Addressee influences on message formulation. Journal of Language and Social Psychology, 6(2), 81-98.

Krauss, R. M., Vivekananthan, P. S., & Weinheimer, S. (1968). Inner speech and" external speech": Characteristics and communication effectiveness of socially and nonsocially encoded messages. Journal of Personality and Social Psychology, 9(4), 295.

Metzing, C., & Brennan, S. E. (2003). When conceptual pacts are broken: Partner-specific effects on the comprehension of referring expressions. Journal of Memory and Language, 49(2), 201-213.

Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. Behavioral and Brain Sciences, 27, 169–226.

R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Schober, M. F., & Clark, H. H. (1989). Understanding by addressees and overhearers. Cognitive Psychology, 21, 211-23

Tomasello, M. (2003). Constructing a language: A usage-based theory of language acquisition.[aHPB].

Wilkes-Gibbs, D., & Clark, H. H. (1992). Coordinating beliefs in conversation. Journal of Memory and Language, 31, 18

BIOGRAPHICAL SKETCH

Ashley Pieper is a Developmental Psychology Ph.D. student at Florida State University. Her advisor is Dr. Michael Kaschak. Ashley's research focuses on social aspects of language, language comprehension and language use. Ashley grew up in Stuart, Florida and attended Florida State University. She graduated cum laude with a Bachelor of Science (B.S.) degree in Criminology, a second major in Psychology, and a minor in Spanish.