

Coordinating Communication in the Wild: The Artwalk Dialogue Corpus of Pedestrian Navigation and Mobile Referential Communication

Kris Liu¹, Jean E. Fox Tree¹, Marilyn A. Walker²

¹University of California, Santa Cruz, Department of Psychology

²University of California, Santa Cruz, Department of Computer Science

E-mail: kylIU@ucsc.edu, foxtree@ucsc.edu, mawalker@ucsc.edu

Abstract

The Artwalk Corpus is a collection of 48 mobile phone conversations between 24 pairs of friends and 24 pairs of strangers performing a novel, naturalistically-situated referential communication task. This task produced dialogues which, on average, are just under 40 minutes. The task requires the identification of public art while walking around and navigating pedestrian routes in the downtown area of Santa Cruz, California. The task involves a Director on the UCSC campus with access to maps providing verbal instructions to a Follower executing the task. The task provides a setting for real-world situated dialogic language and is designed to: (1) elicit entrainment and coordination of referring expressions between the dialogue participants, (2) examine the effect of friendship on dialogue strategies, and (3) examine how the need to complete the task while negotiating myriad, unanticipated events in the real world – such as avoiding cars and other pedestrians – affects linguistic coordination and other dialogue behaviors. Previous work on entrainment and coordinating communication has primarily focused on similar tasks in laboratory settings where there are no interruptions and no need to navigate from one point to another in a complex space. The corpus provides a general resource for studies on how coordinated task-oriented dialogue changes when we move outside the laboratory and into the world. It can also be used for studies of entrainment in dialogue, and the form and style of pedestrian instruction dialogues, as well as the effect of friendship on dialogic behaviors.

Keywords: spontaneous communication, referential communication, naturalistic task, wayfinding, navigation, mobile communication, mediated communication, friends and strangers

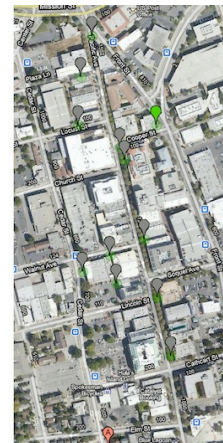
1. Background and Motivation

Previous work on dialogic task-oriented communication has established the fact that a large number of different variables affect the form and linguistic style of the dialogue, such as the personality and cultural identity of the speaker and hearer, their relationship, their shared knowledge of the terrain and their ability to see one another (Anderson et al., 1991, Schober & Clark 1989, Brennan 2005, Mehl et al., 2006, Boyle et al., 1994, Mairesse et al., 2007, Brennan & Clark, 1996).

The Artwalk Corpus is a collection of 48 mobile phone conversations between 24 pairs of friends and 24 pairs of strangers performing a novel, naturalistically situated referential communication task. The task requires the identification of public art while walking around and navigating pedestrian routes in the downtown area of Santa Cruz, California. The task involves two participants: a Director located on the University of California, Santa Cruz (UCSC) campus who is given a map and provides verbal instructions via Skype to a Follower. The Follower must walk around downtown according to the instructions given by the Director and find target objects that the Director is describing. See Fig. 1. The task provides a setting for real-world situated dialogic language and is designed to: (1) elicit entrainment and coordination of referring expressions between the dialogue participants, (2) examine the effect of friendship on dialogue strategies, and (3) examine how the need to complete the task while negotiating

myriad, unanticipated events in the real world – such as avoiding cars and other pedestrians – affects linguistic coordination and other dialogue behaviors.

In labs, researchers usually maintain an environment with relative quiet and minimal uncontrolled distraction while people participate in a specific and circumscribed task. People who are not involved in the experiment do not wander in and out of the testing room and participants do not have to think about anything except for the task (or tasks) at hand. Depending on the population being sampled, participants may have already gone through several similar experiments (for example,



Press SPACEBAR when your partner has found this object

Figure 1. An example of a Director's screen during the task. The map is non-interactive and has grey indicators for potential target locations.

students in a psychological research methods course are often encouraged or required to be research participants).

As a result, participants in lab-based behavioral experiments often experience a task that is already somewhat routine. This routine task takes place within a setting that is more akin to an exam in an isolated booth than it is to the mobile, ambulatory, multi-tasking, and often distracted communication common in everyday life. The Artwalk Task represents a real-world task involving the negotiation of myriad, unanticipated events, while achieving multiple conversational goals. The corpus thus illustrates how dialogue can progress when we are moving around in the world, balancing multiple demands, and engaging in a novel task that still bears some resemblance to an everyday conversation.

The effects of collaboration on language use, and vice-versa, have been assessed with two popular methodologies: the Map Task (Anderson et al., 1991) and the Tangram Task (Clark & Wilkes-Gibbs, 1986; Schober & Clark, 1989). In the Map Task, two participants are given slightly different versions of a map, with one participant playing the role of the Director and the other the Follower. Directors and Followers must then coordinate efforts to reproduce the Director's route onto the Follower's map. This experimental design elicits spontaneous verbal collaboration between the two interlocutors to refer to landmarks and match the route on the Director's map. In the Tangram Task, the Director is shown an array of abstract shapes in a particular order. The Follower is shown a similar or the same set of shapes but they are presented in a different order than the Director. The goal is for the Follower to put the shapes in the same order as the Director. Over successive iterations of the task, the collaborative communicative process can be observed, as interlocutors coordinate on how they refer to the abstract shapes. Like the Map Task, the Tangram Task elicits spontaneous verbal collaboration between the two interlocutors to complete the task: they need to negotiate referents and their referring expressions in order to place them in the same positions as on the Director's template.

In the lab, both Map Task and Tangram Task are abstracted, simplified versions of real-life conversational phenomena: giving and following wayfinding directions that would normally involve moving someone through a complex space becomes two people following a line-drawing path on simple and identical maps. When working with tangrams, negotiating referring expressions potentially gets boiled down into describing component, two-dimensional shapes. Having any two people – regardless of whether they know each other or have anything in common – keep up conversation for at least 30 minutes at the request of an unfamiliar third party becomes nearly impossible without multiple experimenter-assigned prompts that are designed to elicit talk. Additionally, both participants sitting in a bare laboratory room provides little-to-no opportunity to spark conversation that is unrelated to the exact task they are asked to do. Having an experimenter right outside of

the door provides a quick escape hatch when conversational prompts are minimally fulfilled and both participants decide that they are done. In short, dialogue produced in these situations may become perfunctory.

The task developed for the Artwalk Corpus is conceptually similar to the Map Task and Tangram Task, but bears closer resemblance to the original phenomena that both were created to study. It is situated in the world, on a real street, looking for real target objects in a visual field jammed with moving and stationary objects.

Example of Non-Task Distractions on Task Dialogue (Dyad 71)

Follower: Canvassers are everywhere.

Director: Did he hassle you for being on the phone or something?

F: [laugh] Yeah like this guy came up to me and like tried to run in front of me.

D: What?

F: I was just like, "I'm doing a psychology experiment" and he was like, "Oh you're doing the fake phone thing?" and I was like, "What? I'm on the phone." [pause] I could never do that to someone.

In Artwalk, a Director sees a series of public art objects and their locations on a map. The Director is responsible for leading the Follower to each object by giving directions and by describing those objects so that the Follower may take a photo of it. The Follower is not given a map, nor is the Director privy to a dynamically-updated map with the Follower's location. The targets that the Follower has to find were mostly abstract (vs. concrete) artworks. The participants spoke via a Skype-to-mobile phone connection, with the Director using Skype on a computer from the laboratory and the Follower using a mobile phone while wandering the downtown Santa Cruz. Dyads participated in two rounds of the task, finding a single set of five targets twice. In addition to finding and photographing targets, Followers must do what many people do every day: avoid pedestrians and cars, and walk, talk, and look for navigational landmarks at the same time.

Speaking on the phone, with one participant walking on a street, participants are not spared the awkwardness of having to keep up talking to one another. If they are inclined to fill the silence, they have multiple options, such as talking about the task, talking about something that is happening around them, or talking about other topics of joint interest. Each dyad thus completes both Map Task and Tangram Task in the same session in a way that also allows for the elicitation of free-form conversation or small talk.

This real-life route-and-referencing task, where at least one participant is out of the controlled environment, extends existing research by providing a language resource containing examples of how speakers coordinate language and entrain on referring expressions and other linguistic behaviors in the wild. On a street,

participants have much more information to deal with, as well as the competing demands of having to balance logistical concerns of moving through crowded, unpredictable space. In our study, these issues included occlusion of targets and general distraction from task by falling darkness, rain, street performers, cars, and pedestrians, as well as self-presentation issues and object manipulation issues that accompany talking on a cell phone while taking photographs. At the same time, the relatively small neighborhood constrains the number of available paths for the Follower to traverse, which provides some necessary consistency across sessions.

Another difference between our task and other tasks is that communication here is mediated. Mediated communicative settings vary from face-to-face in predictable ways (Fox Tree, Mayer, & Betts, 2011; Clark & Brennan, 1991). Not being in the same physical environment can cause people to talk about their environment when they otherwise would not. Not seeing one's interlocutor can lead people to express or hide information about themselves, and being mobile can introduce topics not found while stationary. Using a device to mediate communication introduces changes to how people communicate, starting from the very beginning of a conversation: for instance, Caller ID can cause changes in openings, allowing people to bypass greetings and engage immediately with the body of the conversation. The loss of cellular signal can cause people to make arrangements in advance about conversational closings ("if the phone dies, that's good-bye") or turn-taking ("if you don't hear from me for a while, I'm in a dead zone but stay on the line").

The cell phone use required by our study also introduces additional grounding constraints (Clark & Brennan, 1991) not present in traditional laboratory tasks. Grounding constraints are factors that affect how communication proceeds under different settings. For example, being able to see one another allows face-to-face addressees to nod for comprehension, but communicating with a phone requires a verbal *uh huh*. These constraints extend beyond those originally discussed by Clark and Brennan (1991) such as *visibility* and *copresence*, and even beyond those discussed by others such as *anonymity* and *locatability* (Hård Af Segerstad & Ljungstrand, 2002) or *multi-tasking awareness* (Fox Tree et al., 2011). One additional grounding constraint mobile phone use introduces might be thought of as the *overhearing constraint*. It may be harder for speakers to design utterances for addressees when they know there are listening bystanders who might be disturbed by the conversation. The presence of bystanders may affect the quality and quantity of talk.

Though it should be noted that we did not collect pairs of friends who then switched off in roles and then paired up with strangers due to the logistical limitations of this task, the even split between friend and stranger dyads (24 each) reflects one of the major factors that went into the collection of the original HCRC Map Task Corpus (Anderson et al., 1991). Relatively few studies

done on the Map Task Corpus report on the differences between friend and stranger dyads, even though it is one of the most straightforward comparisons to make.

There are reasons to believe that people talk differently to friends than they do to strangers. For instance, friends know more about each other than strangers would, which may lead to more effective performance or informal communication. Studies using other corpora and tasks have found differences between the conversations of friends and strangers (e.g., Fox Tree, 2007; Hornstein, 1993; Jucker, Smith, & Lüdge, 2003; Planalp, 1993; Savitsky, Keysar, Epley, Carter, & Swanson, 2011) but only a handful of studies on HCRC Map Task look at these differences: Boyle, Anderson and Newlands (1994) found that friends used more words tokens and turns across the entire dialogue than strangers. Additionally, they found that the mapped paths of friends showed less divergence from each other than the paths of strangers. While Bard, Lickley and Ayelett (2001) found that there was an effect of familiarity on the production of disfluencies, Branigan, Lickley and McKelvie (1999) did not. Trouvain and Truong (2012) found no convergence of laughter for friend dyads, which was expected given Smoski and Bacharoski's (2003) work on another corpus.

The relative paucity of studies that report on the differences between friends and strangers may be due to the fact that the dialogues are quite short (an average of about 10 minutes) and very focused on a straightforward task. As a result, they may be less likely to reflect the characteristics that distinguish the conversations of friends and strangers. By comparison, the average length of the Artwalk Corpus dialogues is just under 40 minutes, contains more complex versions of the Map Task and the Tangram Task, as well as multiple, naturally-occurring points where partners can choose to engage in non-task conversation. If there are substantial differences between how friends and strangers use referential communication in spontaneously-generated speech in naturalistic contexts, they are likely to be found in the Artwalk Corpus.

2. Corpus Description

The corpus consists of transcripts of the dialogue between 24 pairs of strangers and 24 pairs of friends who successfully completed the Artwalk Task. The task involved a Director describing both a piece of art and a route to a Follower, who would photograph the art when found. Successful completion was defined as pairs who found at least 8 of the 10 targets, 4 per each of two rounds. Dialogues ranged from 24 to 55 minutes ($M=39.47s$), with a maximum of 40 minutes spent on the experimental task; the remainder of the time was spent on getting participants orientated at the beginning and back to the starting point at the end. Data collection was carried out in 2012-2013 in Santa Cruz, California.

There are several other pieces of information that are also available to researchers wanting to use the Artwalk Corpus: genders of both participants, whether

the participants are friends or strangers, photos of the targets, a list of coordinates for the starting/ending location and the target objects, a list of selected questions from the post-experiment questionnaire, date and start and finish time of participation, as well as retrospective weather data and sunset times. A list of target order is also provided, since targets were randomized for each dyad, on each round.

Items from the post-experimental questionnaire asked participants to rate themselves on certain aspects of the task: how familiar they were with their partners before the task began; how familiar they are with Santa Cruz; how well they thought they performed; how much effort they put into the task; how comfortable they felt with the task; and how often they spoke on the phone. Participants filled out these questionnaires separately from each other, immediately following the completion of the experiment. It should be noted that though most participants answered most questions, some items are missing for some participants.

Weather information was collected retrospectively from Weather Underground (www.wunderground.com). This included information on temperature, humidity, wind and presence of clouds. While Santa Cruz has a moderate climate year-round, there was the possibility of inclement weather would change Followers' behavior (which would then also influence the Director's behavior, even though the Director was located in the lab the entire time). Similarly, participants may be more willing to rush through a task if they were at the end of their day or if dusk made it harder to distinguish darker targets against dark and busy backgrounds, or in smaller alleys. It is also notable that crowds fluctuate depending on weather, time of year and time of day. People spend more time walking around outside when the weather is warm (de Montigny, Ling, & Zacharias, 2012). Santa Cruz, a college town and summer tourist destination, tends to be more crowded when school is in session and in the height of summer. Like many other cities, there are morning and evening rush hours and times when people are more likely to venture out for meals or to run errands. Artwalk is a task that can vary in difficulty depending on a number of external factors. The corpus and additional data are available at nlds.soe.ucsc.edu/corpora.

2.1 Participants

Participants were University of California, Santa Cruz undergraduates who participated for either course credit or a \$10 gift card. Participants are either assigned to perform the task downtown (Follower, on a cell phone) or in a campus lab (Director, through Skype). Friend dyads, which were difficult to recruit through the participant pool, were largely recruited through word-of-mouth and on-campus ads. Stranger dyads mostly came from the participant pool.

2.2 The Artwalk Task

Downtown Santa Cruz is filled with public art such as sculptures, mosaics, and murals. Ten target art pieces

were identified based on their potential to elicit a wide variety of descriptive words. All were located within a two-by-six block area. Most targets were close to other artworks that often needed to be explicitly eliminated by the Follower as a potential target. During the year of data collection, three pieces were removed by the city; we replaced these targets with new targets that were geographically close to the original target.

Example of Wayfinding Leading Into Target Finding (Dyad 27)

D: Okay um alright so right where you found that um painting
 F: *Uh huh*
 D: *You want* to keep going on Walnut until you see cedar street
 F: Okay
 D: So keep going um down on *Walnut* yeah
 F: *Walnut* alright...alright I'm about to cross Cedar
 D: Uh okay wait *from Walnut you sh- yo-* from Walnut you should uh make a left on Cedar
 F: *[unintelligible]* O:kay
 D: Okay now *don't* go to- don't go too far cause you're going to see a uh it looks like a like a painted brick
 F: *Uh huh* a painted brick
 D: Yeah it has blue and then like some lines on it
 F: Like grayish lines around it?
 D: [tapping] Hello?
 F: Like it's outlined in a grey-
 D: It- kind of it it looks like blue painted with stripes on it
 F: With what on it?
 D: Uh stripes or lines
 F: Blue painted with stripes h
 F: [laugh] there's this giant mural on the side of Petroglyph
 D: Uh huh
 F: And there are lots of bright colors and squares of things
 D: Do you see a square *with* with uh any blue in it?
 F: *I'm-*... um there are blue squares yes
 D: Okay does any of them have this weird like design that has like lines on it looks like Ts two Ts *together*
 F: *Oh* like does it kinda have like a zipper?
 D: Yeah exactly
 F: Ah alright I found one [laugh]

Directors were shown a single target object at a time along with a map with the target's location highlighted (see Figure 1). Target order was randomized and each target was given an 8-minute time limit. If participants had not proceeded to the next target by hitting a key to indicate they found the target, they were automatically moved to the next one and were obligated to stop their search for the timed-out target. After all 5 targets were found once (Round 1), targets were

re-randomized and presented again for the pairs to find (Round 2). Participants were instructed to not deviate from the order set for them; the handful of pairs that did were excluded from this study. Followers were instructed to take pictures of the targets as they located them, which were then checked by the experimenters after the experimental session was concluded.

Similar to the original Tangram Task, Directors had no access to the targets identified by Followers as they could not see the photos, and so had to trust that the Follower had found the correct target.

Example of Task/Non-Task Blended Dialogue (Dyad 28)

F: Ew: a bird pooped on me [laughing]
 D: Oh my god really *ew:*

F: *Ye:s* [laughing]
 D Oh my god where did it get you?
 F: So: disgusting on my sleeve e:w
 D: Oh my god
 F: And I'm wearing a white shirt
 D: Oh my god I'm so sorry
 F: [laughing] It's okay... Oh my god [unintelligible]
 D: I can't believe that just happened, I'm so sorry [both laughing]
 F: Okay, Okay I'm taking a picture of the [unintelligible; previous target] [laughing]
 D: [laughing] I'm so sorry
 F: Oh cra:zy
 D: Okay and then you're gonna want to go back to where the tiles were between Lincoln and Walnut
 F: Okay gosh. This is so embarrassing
 D: I can't believe that actually happened
 F: [laughing] [unintelligible] Oh so nasty
 D: Ew:, oh my god
 F: Gosh darn it, I'm checking myself out in a mirror making sure I look okay, don't wanna be walking around [unintelligible]
 D: Okay
 F: Okay I'm going back to the tiles
 D: Okay
 F: Do you know how much time I have?
 D: Uh I we're doing really well. Actually I think this is the last one and you still have like 12 minutes or something.

2.3 Corpus Creation and Characteristics

Dialogue was manually transcribed by research assistants using ExpressScribe and checked a second time by a different research assistant.

This complex corpus contains dialogue typical for referential communication tasks, but because the task is designed to be a combination of Tangram and Map tasks performed in a real-world setting, the dialogues are melded in a way that is not found in laboratory-based tasks. There are instances of poor cellular reception or noisy surroundings causing participants to check whether their partners are still with them, as well as subdialogues

discussing the distractions that occur when on a city street. Also, because of the relatively long delay between targets, some participant pairs engaged in casual side conversations about task-relevant and task-irrelevant topics. Small talk, side commentary about the study, joking and conversations about other topics were present across the corpus – and often blended in various combinations – presenting a fuller picture of how goal-oriented conversations unfold when not in a strictly controlled environment. Not all dialogues had examples of every possible type of conversation. At the same time, few dialogues that contained task-irrelevant dialogue showed a perfect separation of different task-related and task-unrelated conversational threads.

The corpus includes examples of many spontaneously produced phenomena of interest to researchers. These include phenomena such as repetitions, false starts, and repairs (Fox Tree, 1995), fillers such as *um* and *uh* (Clark & Fox Tree, 2002; Fox Tree, 2001, 2002), and hedges such as *sort of* or *I think* (Liu & Fox Tree, 2012). They also include discourse markers like *oh*, used to indicate upcoming disjuncture in speech or to express attitudes (Fox Tree & Schrock, 1999), *like* and *you know*, used to indicate that information is being left unexpressed (Fox Tree, 2006, 2007; Fox Tree & Schrock, 2002), and other discourse markers such as *actually* and *really* (Fox Tree, 2000). The corpus also includes many examples of other devices that coordinate across conversational participants, e.g. backchannels like *uh huh* (Tolins & Fox Tree, 2014, 2015) and agreement markers such as *right* and *okay* (Bangerter & Clark, 2003).

Other participant and experimental session data, such as responses to post-experiment questionnaire items, date/time of participation, gender breakdown, friend or stranger dyad, sunset times and weather data were recorded in a spreadsheet. Photos and coordinates of the targets are in a separate document.

3. Sample Data Analysis

As an example of the kind of analyses that can be conducted with this corpus, we present here some of the discoveries we have made (see Liu, Fox Tree, Blackwell, & Walker, under review, for more details). One major discovery was that entrainment can be found in this naturalistic task as it has been in laboratory tasks.

We analyzed the number of Director Descriptors used between Round 1 and Round 2. Descriptors were defined as unique-within-round adjectives or nouns that were descriptive of the target artwork, such as colors, shapes, media type (e.g., painting, sculpture, metal, stone) and patterns (e.g. striped). Descriptors were counted for each piece of target artwork, for each round separately. Directors used fewer Descriptors in Round 2 ($M = 4.17$, $SD = 1.77$) than in Round 1 ($M = 12.23$, $SD = 3.75$), $t(47) = 14.06$, $p < .0001$, 95% CI [6.91, 9.21], $d = 2.03$ (Liu et al., under review). Thus, on this measure, we replicated effects found in the laboratory “in the wild”: Participant

pairs became more efficient at referring to their targets in Round 2 than they had been in Round 1.

While we tested the presence of entrainment using the number of descriptive words used (Liu et al. under review), there are many potential analyses that may be undertaken in the future. For example, researchers may test efficiency based on whether descriptions were holistic or piecemeal, abstract or concrete, co-occurring with directional information or not. Researchers may also explore whether repetition of words across conversational participants enhanced or hindered efficiency (Brennan & Clark 1996). There are also a number of automatic analyses of entrainment or alignment phenomena in terms of structural priming or referring expression content selection that could be carried out using this corpus (Reitter et al, 2006, Reitter & Moore, 2014, Gupta & Stent, 2005, Jordan & Walker, 2005, Guhe & Bard 2008).

Example of Attempting to Use Expertise on Santa Cruz and Common Ground (Dyad 16)

F: There's a roller coaster here, but it's got like, yeah it looks like a prison cell because of the way the shot's like drawn... but it's not that it's not as small as you described it

D: How tall is it? or how big is it?

F: U:h across, I'm gonna say like 5 feet, 6 feet

D: Maybe it's- okay maybe it's bigger than what I'm describing it to be, but yeah it kinda looks like train tracks, and it kinda looks like a ladder like tilted, like on the floor

F: Yeah

D: Not tilted

F: Yeah, it's a- cause the way the roller coaster, have you seen the Giant Dipper?

D: The Giant Dipper?

F: Are you seen- like have you been to the boardwalk

D: No. I mean I have been there but I don't know what you're talking about

We have also used insights and sample utterances from Artwalk for experiments aimed at developing models of adaptive language generation and entrainment in pedestrian direction-giving dialogues (e.g., Hu, Halberg, Jimenez, & Walker, 2014). We show that hedges and particular types of entrainment are perceived as more friendly than dialogue utterances modeled on existing direction giving systems, such as Google Maps.

There are also opportunities for new explorations across dyad dynamics. For example, dyads with more small talk may be more or less efficient than those with less small talk. Dyads with more directional problems may also be those who take longer to identify the art, or the two could be unrelated. Art that has more similar descriptions reached about it across dyads may be identified more quickly than art that has more

idiosyncratic descriptions. Friends may be more likely to adopt one method of describing art, and strangers another. Dyads who are both familiar with Santa Cruz may have an easier time negotiating referents based on shared common ground (i.e., expertise with the neighborhood layout or shops), as well as an easier time with wayfinding.

In addition to testing the effect of the number of descriptive words used on entrainment, we also conducted analyses based on acquaintanceship (friends or strangers). We found that friends and strangers were equally efficient in their referring expression negotiation (Liu et al, under review). We also found some evidence for adaptation across friends, but not strangers. Directors in friend dyads altered their behavior in the second round of the task based on how many turns it took them and their Followers to identify a target artwork in the first round. The more turns friend dyads took in Round 1 to identify the artwork, the more descriptors the friend Directors used in Round 2. Stranger dyads did not adapt in this way.

The additional data included may be interesting to those wishing to explore, or take into account, how other factors besides those controlled by the task can influence communicative behavior in wayfinding, referring expression negotiation, or small talk. For instance, preliminary correlational analyses can suggest some more fruitful and less fruitful analytical paths that may be taken: the effect of gender pairing (female Director/female Follower, female Director/male Follower, etc.) on number of Descriptors used was virtually nil. This null finding, however, may be due to the imbalanced percentages (about 40% of the corpus are female/female pairs, while mixed gender accounts for 30%). In comparison, weather may be an important covariate, particularly for the Follower: warmer weather was associated with increased use of Descriptors by Followers in Round 1, $r(46) = .35, p = .02$. Gustier conditions led to greater Descriptor use by Followers in Round 2, $r(46) = .37, p = .01$. The ambient conditions of a quiet laboratory room and of a city street subject conversations to different pressures.

4. Conclusion

The Artwalk Task provides a method for collecting information both on referential communication and also on navigation. While people do not generally get led around streets in search of public art, the task does approximate the route-finding and location-identification that is part of more common activities, such as looking for stores, restaurants, or more famous landmarks. The structure of the task allowed for conversations to flow more freely than laboratory-based tasks and allowed for a wider array of dialogue from participants: some participants stayed strictly on task and others had somewhat meandering conversations while the Follower was walking from one target to the next. The fact that the dialogue was created alongside a mobile task introduced numerous opportunities for small talk and emotional

expression, as well as multiple levels of coordination, including coordinating on identifying art, navigating streets, avoiding environmental obstacles, and carrying on a conversation.

The main features of the Artwalk Corpus is that it includes including relatively long dialogues (25-60 minutes of nearly continuous talk), two different referential tasks (Map Task and Tangram Task), task-oriented and casual conversation, and an even split between friend and stranger dyads. The corpus has numerous examples of spontaneous communicative phenomena including speech disfluencies, fillers, discourse markers, and backchannels. It also includes expressions of genuine, spontaneously-produced emotion such as frustration and surprise, as well as other social phenomena such as apologies and humor.

The Artwalk Corpus opens the door for exploration of a number of linguistic and pragmatic phenomena. For example, researchers may be interested in strategies that are used when dyads switch between different modes of talk, such as switching from small talk to wayfinding, switching from navigation to object-identification, or switching from a period of silence to talking again. Researchers may also be interested in recovery strategies, such as how dyads identify that they have miscommunicated and how they resolve that miscommunication. The corpus provides an opportunity to explore how people adapt to their partner's conversational styles (such as Linguistic Style Matching cf. Niederhoffer & Pennebaker, 2002) or task approaches, over the course of the conversation.

This corpus also provides a window into mobile communication, as many people engage in goal-oriented conversations while walking down streets and navigating relatively crowded public spaces; the fact that one conversational partner is connected via Skype on a computer (similar to being limited to a single location via landline) and the other connected via mobile phone might be particularly useful as it is not a particularly common feature of many corpora. Those studying navigation may be interested in how people describe and coordinate directions or distances. Researchers studying emotions may be interested in how people defuse tension that arises from being frustrated by the task. Researchers developing models of adaptive virtual

agents may be interested in the variety of adaptations displayed in Artwalk (e.g. Hu, Walker, Neff, & Fox Tree, 2015; Tolins, Liu, Wang, Fox Tree, Neff, & Walker, 2013, 2016, Hu, Dick, Chang, Bowden, Neff, Fox Tree, & Walker, 2016). In short, Artwalk is a corpus rich in analytical possibilities.

5. Acknowledgments

This work was supported by NSF Grant IIS-1044693. We thank Natalia Blackwell for her contributions to this project. We also thank our RAs: Jennifer Fong, Steven Ethington, Erin Hiscock-Wagner, Beth Hart, Tyler Pilgrim, Haley Bieseimer, Megan Kostecka, Natalie Serourian, Jason Tharp, Sean Tang, Steven Finet, Lisa Tatoian, and Melissa Kitzmiller.

6. References

- Anderson, A., Bader, M., Bard, E., Boyle, E., Doherty, G. M., Garrod, S., & Weinart, R. (1991). The HCRC Map Task Corpus. *Language and Speech, 34*, 351-366.
- Bangerter, Adrian, Clark, Herbert H., 2003. Navigating joint projects with dialogue. *Cognitive Science, 27*, 195-225.
- Bard, E. G., Lickley, R. J., & Aylett, M. P. (2001). Is disfluency just difficulty? In *Disfluency in Spontaneous Speech (DiSS'01)*, ISCA Tutorial and Research Workshop (ITRW), 97-100.
- Boyle, E. A., Anderson, A. H., & Newlands, A. (1994). The effects of visibility on dialogue and performance in a cooperative problem solving task. *Language and Speech, 37*(1), 1-20.
- Branigan, H. P., Lickley, R. J., & McKelvie, D. (1999). Non-linguistic influences on rates of disfluency in spontaneous speech. In *Proceedings of the XIVth International Congress of Phonetic Sciences*. San Francisco.
- Brennan, S. E. (2005). How conversation is shaped by visual and spoken evidence. In *Approaches to studying world-situated language use: Bridging the language-as-product and language-as-action traditions*, 95-129. Trueswell and Tanenhaus, Eds.
- Brennan S.E. and Clark, H. H (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning, Memory and Cognition, 22*(6)
- Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. In J. M. L. L.B. Resnick, & S.D. Teasley (Ed.), *Perspectives on socially shared cognition*. Washington, DC: APA Books.
- Clark, H. H., & Fox Tree, J. E. (2002) Using uh and um in spontaneous speaking. *Cognition, 84*, 73-111.
- Clark, H. H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition, 22*, 1-39.
- de Montigny, L., Ling, R., & Zacharias, J. (2011). The effects of weather on walking rates in nine cities. *Environment and Behavior, 44*(6), 821-840.
- Fox Tree, J. E. (1995). The effects of false starts and repetitions on the processing of subsequent

Example of Flirting (i.e., Naturally-Occurring Conversational Phenomena) (Dyad 26)

- D: Yeah so tell I'll stay on until you get back there, cuz after this I'm never gonna talk to you again and that will be a shame
- F: Wow it sounds like you're hitting on me sir
- D: Well I don't want it to come off like that but I'm gonna miss you
- F: We uh we should probably hang out after this you sound interesting
- D: Yeah fine where do you live?

- words in spontaneous speech. *Journal of Memory and Language*, 34, 709-738.
- Fox Tree, J. E. (2000). Coordinating spontaneous talk. In L. R. Wheeldon (Ed.), *Aspects of Language Production* (pp. 375-406). Philadelphia: Psychology Press.
- Fox Tree, J. E. (2001). Listeners' uses of um and uh in speech comprehension. *Memory and Cognition*, 29(2), 320-326.
- Fox Tree, J. E. (2002). Interpreting pauses and ums at turn exchanges. *Discourse Processes*, 34(1), 37-55.
- Fox Tree, J. E. (2006). Placing *like* in telling stories. *Discourse Studies*, 8(6), 749-770.
- Fox Tree, J. E. (2007). Folk notions of *um* and *uh*, *you know*, and *like*. *Text & Talk*, 27-3, 297-314.
- Fox Tree, J. E., & Schrock, J. C. (1999). Discourse markers in spontaneous speech: Oh what a difference an oh makes. *Journal of Memory and Language*, 40(2), 280-295.
- Fox Tree, J. E., & Schrock, J. C. (2002). Basic meanings of you know and I mean. *Journal of Pragmatics*, 34, 727-747.
- Fox Tree, J. E., Mayer, S. A., & Betts, T. E. (2011). Grounding in instant messaging. *Journal of Educational Computing Research*, 45(4), 455-475.
- Guhe, M., & Bard, E. G. (2008). Adapting referring expressions to the task environment. In *Proc. of the 30th Annual Conf. of the Cognitive Science Society (CogSci)* (pp. 2404-2409).
- Gupta, S., & Stent, A. (2005). Automatic evaluation of referring expression generation using corpora. In *Proc. 1st Workshop on Using Corpora in NLG*.
- Liu, K., Fox Tree, J. E., Blackwell, N., & Walker, M. A. (under review) Referential communication between friends and strangers in the wild.
- Liu, K. & Fox Tree, J. E. (2012). Hedges enhance memory but inhibit retelling. *Psychonomic Bulletin & Review*, 19(5), 892-898.
- Hård Af Segerstad, Y., & Ljungstrand, P. (2002). Instant messaging with WebWho. *International Journal of Human-Computer Studies*, 56(1), 147-171.
- Hornstein, G. A. (1985). Intimacy in conversational style as a function of the degree of closeness between members of a dyad. *Journal of Personality and Social Psychology*, 49, 671-681.
- Hu, C., Walker, M. A., Neff, M., & Fox Tree, J. E. (2015). Storytelling agents with personality and adaptivity. *Intelligent Virtual Agents* (pp. 181-193).
- Hu, Z., Dick, M., Chang, C.N., Bowden, K., Neff, M., Fox Tree, J. E., Walker, M. (2016). A corpus of gesture-annotated dialogues for monologue-to-dialogue generation from personal narratives. *Proc. Language Resources and Evaluation Conf. (LREC)*.
- Hu, Z., Halberg, G., Jimenez, C., & Walker, M. (2014). Entrainment in pedestrian direction giving: How many kinds of entrainment. *Proc. IWSDS 2014*, 90-101.
- Jordan, P.W., Walker, M.A. (2005). Learning Content Selection rules for Generating Object Description in Dialogue, *J. Artif. Intell. Res. (JAIR)* 24, 157-194.
- Jucker, A. H., Smith, S. W., & Lüdge, T. (2003). Interactive aspects of vagueness in conversation. *Journal of Pragmatics*, 35, 1737-1769.
- Mairesse, F., Walker, M. A., Mehl, M. R., & Moore, R. K. (2007). Using linguistic cues for the automatic recognition of personality in conversation and text. *Journal of artificial intelligence research*, 457-500.
- Mehl, M. R., Gosling, S. D., & Pennebaker, J. W. (2006). Personality in its natural habitat: manifestations and implicit folk theories of personality in daily life. *Journal of and social psychology*, 90(5), 862.
- Niederhoffer, K., & Pennebaker, J. (2002). Linguistic style matching in social interaction. *Journal of Language and Social Psychology*, 21(4), 337-360.
- Planalp, S. (1993). Friends' and Acquaintances' Conversations II: Coded Differences. *Journal of Social and Personal Relationships*, 10, 339-354.
- Reitter, D., Keller, F., & Moore, J. D. (2006). Computational modelling of structural priming in dialogue. In *Proc. of the Human Language Technology Conf. of the NAACL* (pp 121-124).
- Reitter, D., & Moore, J. D. (2014). Alignment and task success in spoken dialogue. *Journal of Memory and Language*, 76, 29-46.
- Savitsky, K., Keysar, B., Epley, N., Carter, T., & Swanson, A. (2011). The closeness-communication bias: Increased egocentrism among friends versus strangers. *Journal of Experimental Social Psychology*, 47, 269-273.
- Schober, M. F., & Clark, H. H. (1989). Understanding by addressees and overhearers. *Cognitive Psychology*, 21, 211-232.
- Smoski, M., & Bachoroski, J.-A. (2003). Antiphonal laughter between friends and strangers. *Cognition & Emotion*, 17, 327-340.
- Tolins, J. & Fox Tree, J. E. (2014). Addressee backchannels steer narrative development. *Journal of Pragmatics*, 70, 152-164.
- Tolins, J. & Fox Tree, J. E. (2015). Overhearers use addressee backchannels in dialogue comprehension. *Cognitive Science*, 1-23.
- Tolins, J., Liu, K., Wang, Y., Fox Tree, J. E., Walker, M.A. Neff, M. (2013). Gestural adaptation in extravert-introvert pairs and implications for IVAs. In *International Conf. on Intelligent Virtual Agents, (LNCS/LNAI) 8108* (pp. 481-482).
- Tolins, J., Liu, K., Wang, Y., Fox Tree, J. E., Walker, M. Neff, M. (2016). A multimodal corpus of matched and mismatched extravert-introvert conversational pairs. In: *Proc. the Language Resources and Evaluation Conference (LREC)*.
- Truong, K. P., & Trouvain, J. (2012). Laughter annotations in conversational speech corpora-possibilities and limitations for phonetic analysis. *Proc. of the 4th Int. Workshop on Corpora for Research on Emotion Sentiment and Social Signals*, 20-24.