UTEP Technical Report UTEP-CS-21-23

The UTEP Corpus of Dissatisfaction in Spoken Dialog

Jonathan E. Avila, Nigel G. Ward, Aaron Alarcon

jonathan.edav@gmail.com, nigelward@acm.org, aaronalarcon2368@gmail.com

Department of Computer Science
University of Texas at El Paso
500 West University Avenue
El Paso, TX 79968-0518

March 2, 2021

We present a corpus of spoken dialogs collected to support research in the automatic detection of times of dissatisfaction. We collected 191 mock customer-merchant dialogs in two conditions: one where the scripts guided the participants to a satisfactory, mutually agreeable outcome, and one where agreement was precluded. Most dialogs were 1 to 5 minutes in length. The corpus and metadata are freely available for research purposes.

keywords: user satisfaction, commercial dialogs, stereo recordings

1 Motivations

The general motivation for this data collection is to support the development of automated methods for determining when, in dialog, the user is dissatisfied. These automated methods could be useful, for example, for generating data that could be used to tune dialog systems, so that they learn to avoid behaviors that cause dissatisfaction.

The immediate practical motivation is the desire to automatically detect abusive merchants from their telephone behavior. One of Google's services is connecting customers with merchants who advertise phone numbers for specific services. The automatic identification of abusive advertisers, who habitually provide inadequate services or attempt to scam callers, would support withholding guarantees or blocking advertisements.

No existing, available corpus is adequate for these purposes. Corpora of commercial dialogs exist but none appear to be generally available. In particular, call-center dialogs have been used to study related problems [Kim, 2008, Ang et al., 2002, Li et al., 2019, Irastorza and Torres, 2018], but again these are not generally available. Many corpus-based studies of the spoken language expressions of emotions and stances have been done, but expression of disappointment has not

been a focus of previous research and no suitable resources appear to be available. We therefore collected our own corpus.

2 Scenarios

We wanted to collect data matched across the two conditions: one where the participants reached a satisfactory, mutually agreeable outcome, and one where they did not. In addition, we wanted to have data from each participant in both conditions.

We would have liked to collect naturally occurring dialogs. However, due to privacy issues and our need to control the conditions, we collected acted conversations, with one person playing the customer role and another playing the merchant role.

Nevertheless, we worked to make the dialogs as realistic as possible. Each participant was instructed to call as if they wanted some specific product of service according to descriptions that we provided. Products or services were chosen to be things that most people could imagine wanting, for example an appointment to get a dog washed or a tire repaired. Wanting participants to speak using their own words, we did not provide scripts, but instead just short paragraphs explaining the scenario: what they wanted, how much they could pay, and other details such as their car model or the size of their dog, depending on the purchase. As seen in Appendix A, many of the scenarios were under-specified and participants were free to elaborate. We observed that participants had no difficulty assuming the assigned role, and several were quite enthusiastic and convincing as they followed the scenarios.

The merchants were played by confederates. Confederates required training, plus the ability to follow scenarios and enact the specified condition. There were two possible conditions for each scenario: in the "legit merchant" condition, instructions guided the confederate to provide the desired outcome at a cost within the customer's specified budget. In the "scam merchant" condition, the confederate was constrained to offer only an inferior or unwanted product or service and/or an unreasonable price. In this condition, the customer was therefore constrained to either accept the bad deal, or to decline the offer; in either case, failing to achieve the goal specified for them.

Of course, no money actually changed hands, and the participants knew that the scenarios were not real. Nevertheless, we wanted them to take the task seriously. To encourage this we provided an incentive: a bonus for every successful transaction, that is, each one completed within the assigned budget.

For each participant, only about half of the calls were in the "legit" condition, and thus only about half presented the opportunity to earn the bonus. To avoid disappointing the participants, they were told in advance that earning the bonus would not always be possible, but we did not provide further details, as we were concerned that participants might become suspicious and obsess about identifying the point at which the dialog was clearly going to fail. The instructions and consent form are included as Appendices B and C. Our research goals were revealed to the participants in the final debriefing, as seen in Appendix D.

3 Participants and Confederates

All data was collected through crowdsourcing. Originally we planned to record on campus; however, the COVID-19 pandemic forced us to move the collection online.

We needed a crowdsourcing platform that allowed us to record telephone conversations, and found that Prolific was able to support this. Asking participants to make a phone call potentially exposes private information, namely their phone number and/or name. Prolific agreed that the study would be acceptable if we took precautions to protect participants' identities, such as never recording incoming phone numbers. The confederates took the calls on their own phones, but their privacy was protected by routing calls through Google Voice.

The Prolific platform enabled us to set criteria for the participants. We selected for workers who indicated that they were fluent in English, had a phone, and were comfortable taking part in a study that may contain deception. We found that most participants called from outside the US, probably because the crowdsourcing platform is based outside US, and most appeared to be non-native speakers. Most calls from outside the US came from Portugal or Poland. Participants earned $5.00 or more in base pay and up to $2.00 in bonus pay for approximately 20 minutes of their time.

The confederates included two of the co-authors, other members of the research lab, and other UTEP students. Confederates were selected based on their people skills and ability to creatively follow scripts. They were briefly trained by the research team in how to handle customer calls following the scenarios and conditions, and underwent human-subjects training before interfacing directly with participants, as detailed in Appendix E.

4 Recording Conditions and Equipment

To support acoustic analysis, we needed a recording setup capable of recording the interlocutors on separate tracks. We considered various call recording applications available for Android or iOS, but none appeared to support stereo recording. We also considered voice chat applications, but these all would have required participants to sign in or create an account. We therefore chose a hardware solution.

The hardware setup is composed of a phone, microphone, headphones, and a computer, running Audacity, as the recording device. Figure 1 shows the components and connections. Critically, the computer's microphone input is a stereo signal with the customer's voice in one channel and the merchant's in the other. Due to the COVID-19 pandemic, we could not use the lab, so we had each confederate replicate the setup using their own phone and computer, in their own room. We provided the cables and, if needed, a headset. Remote debugging of the setups was a challenge, but, in the end, seven of the confederates were able to get it working. We neglected to specify the intended channel assignments for component 7, namely phone output (customer) in the left channel and microphone input (merchant) in the right channel, so for some confederates' recordings we had to correct this in post-processing.

Each confederate was thus responsible for starting the call recording, participating in the merchant role, collecting the participant's ID, ending the call recording and saving the file, and entering notes in a "call log" spreadsheet. These notes included an indication of whether the

call was intended to be a "scam" call, that is whether the confederate was following a script designed to prevent them from satisfying the customer's need, and whether the customer showed dissatisfaction or annoyance during the call.

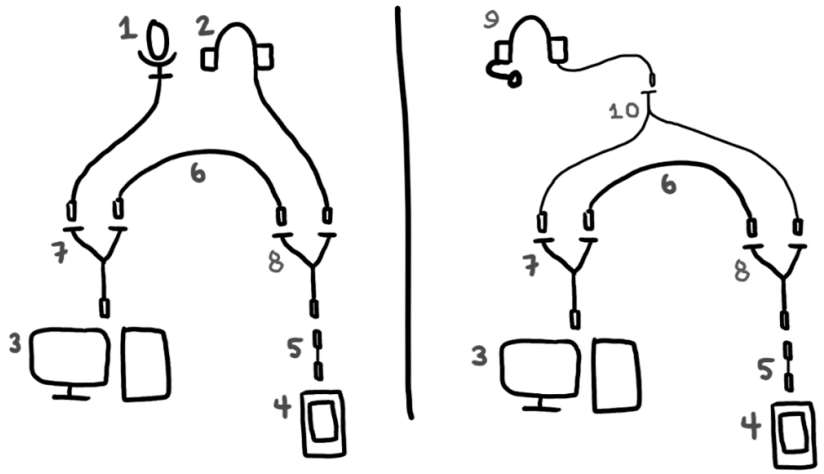

Figure 1: Recording setups: separate microphone and headphones on left; integrated headset on right. All plugs and jacks are 3.5mm. (1) microphone; (2) headphones; (3) computer with mic-in jack; (4) phone with audio jack; (5) Lightning to headphone jack adapter (for iPhone without audio jack) or USB-C to headphone jack adapter (for non-iPhone without audio jack); (6) male to male stereo cable; (7) male to dual female stereo breakout cable; (8) stereo male to dual stereo female Y-cable; (9) headset; (10) female to dual male headset splitter.

We needed to collect a participant identifier for each call so that we could pay participants appropriately. Thus the dialogs were generally preceded or followed by an interaction in which we obtained either their 16-character Prolific ID or a nickname linked to that ID, or gave them a completion code. Although present in the recordings, these interactions should generally be excluded from analysis.

5 Data Collection

Most of the calls were recorded over 16 sessions, held one per day, on various dates in December 2020 through February 2021. We had 67 participants call in. We originally hoped that each participant would interact with multiple confederates, in both conditions and in various scenarios, but the availability of confederates was a limiting factor, and for many participants we were able to obtain only two dialogs.

Ideally all dialogs would have been gathered using a consistent protocol, but in fact we made small adjustments as we went, to address four problems. 1) Many potential participants clicked on our task in Prolific, but did not complete the work, i.e. did not make any calls. We clarified and streamlined the instructions. We also abandoned book-ahead scheduling, in favor of setting aside blocks of time for recording sessions, typically about 2 hours, arranging for all available confederates to be at their phones during the whole period, and posting the task in Prolific just before the start of the session. 2) We still had an overall low sign-up rate, so we increased the

base pay from \$5.00 to \$7.50 and scheduled more sessions for mornings, to be more convenient for callers from Europe. 3) Some of our confederates began to lag in enthusiasm, understandably, because we had them mostly just standing by, idly waiting for callers. We addressed this by raising their compensation also, retroactively, to \$100 for completing their training and debugging the setup, \$10 for each hour spent available to take calls, plus \$1 for each call handled. 4) Many participants accepted even ridiculously bad offers, without showing disappointment or even trying to ask for a better deal. While the protocol already included a bonus opportunity to incentivize participants to follow the instructions and decline offers higher than the target price, we decided to tighten the screening criteria, to require the participants to have in Prolific a higher minimum approval rate and higher number of previously submitted responses. Overall we feel that the these adjustments did not significantly change the nature of the dialogs we obtained.

We received a total of 202 calls. Of course more would have been better, but this is what our time and resources permitted.

6 Post-Processing

In post-processing we matched the records in the *call-log* spreadsheet with the recordings and categorized each. For two of these, the participant did not complete the final debriefing, and thus could not be deemed to have given informed consent. These recordings were therefore discarded.

There were 16 calls from repeat participants, who completed one session and then returned for another. In these repeat calls the participant had already completed debriefing and was no longer naive to the exact aim of our study. These dialogs are tagged *non-naive* and should probably be excluded from analyses, at least for evaluation purposes. These are still, however kept in the corpus, as aspects of these dialogs may still be informative. Some researchers may also choose to use them in their entirety, especially if they are skeptical about whether participants carefully read debriefing information, or about whether such information would likely affect behavior.

There were in addition a few calls where either the merchant or the customer departed from the instructions. We identified these calls from the indications noted by the confederates, which we subsequently verified. In the *mixed_n_y* category, the customer gave up before completing the task or was noticeably dissatisfied, even though the confederate was intending to provide a satisfactory outcome. This happened, for example, when a participant misunderstood something the merchant said. In the *mixed_y_n* category, the confederate was set to offer only an unsatisfactory outcome, but agreement was nevertheless reached. Sometimes this was the participant's fault, when they happily accepted a bad offer, deviating from their instructions, and sometimes the confederate's fault, when they gave in to the customer's counter offer or entreaties, deviating from their instructions. For our purposes such behaviors were frustrating, but they are also realistic: the participants and confederates apparently felt a social imperative to be cooperative and reach agreement, even though they were total strangers and were given instructions intended to prevent this.

Yet another unexpected category was *doomed_2*. In these dialogs, the scenario was such that the merchant and customer should not have reached agreement, and the customer was aware that their goal was being thwarted, but in the end gave in anyway and accepted a bad offer. Figure 3 illustrates.

One of the planned-for categories was *doomed_1*. In these dialogs the instructions were designed to prevent reaching agreement, and both customer and merchant held fast to their assigned constraints, and concluded the dialog without reaching agreement. Figure 2 illustrates.

The other planned-for category was *successful*. In these dialogs the merchant and customer were set to reach a mutually satisfactory outcome and in fact do. It is not the case that these dialogs are entirely smooth sailing, as some of our confederates, getting into their roles, threw in some upselling attempts, but these were successfully declined, and the conversations ended agreeably. Figure 4 illustrates.

Table 1 gives the counts for each category.

Label	Description	Count
non-naive	customer was not naive to the aims of our study	16
mixed_n_y	customer unexpectedly abandoned an on-track dialog	6
mixed_y_n	either customer or merchant deviated from instructions, in order to salvage a set-for-failure dialog	54
doomed_2	customer gave in and accepted a bad offer	27
doomed_1	merchant and customer failed to reach agreement, as expected	25
successful	merchant and customer were set to reach a mutually satisfactory outcome, and succeeded	63
Total		191

Table 1: Category descriptions and counts.

7 Observations

To enable comparisons, we wanted to record each participant interacting with multiple confederates in different conditions. However, we did not achieve a balanced collection. Calls that were set to fail went off-track more often than calls that were set to succeed (hence more dialogs labeled as *mixed_y_n* than *mixed_n_y*).

Some participants were more persistent than others. This is most noticeable in calls set up for the merchant and customer to fail to reach agreement, where participants tried to get the price they wanted by negotiating with the merchant, suggesting ways of "meeting in the middle," by means not offered in the script.

8 Suggested Uses

For the Google scenario, where we look for abusive merchants versus honorable ones, we compare *doomed_1* and *doomed_2* dialogs versus *successful* dialogs. To prioritize evaluation for this purpose, we reserved 17 from *doomed_1* and 21 from *doomed_2*, plus 38 from *successful*, as the evaluation set (test set). Of the remaining files, we split the training and development sets as 4 from *doomed_1*, 3 from *doomed_2*, and 12 from *successful* for each set, as seen in Table 2.

...
2:10 M How can I help you today?
2:12 C Well, I have a Honda Civic and I need to repair a tire that is flat.
2:22 M Alright, you got a flat? So right now our shop's pretty busy and so if you wanted it repaired today we're gonna have to add a forty dollars just for convenience because we're really booked today and then it would be a ten dollar tire repair. But, I could help you out with a deal. I can give you a bundle and I can waive that convince fee. So let me tell you some bundles we have.
2:45 C Alright.
2:46 M So the first one we have is the "Dream Car" bundle. It comes with a car detail, a tire rotation, a full tire inspection, and the tire repair for only two hundred ten dollars.
2:57 C Alright, it's off my budget.
3:01 M Little bit off your budget? How about the "Premium" bundle then? It comes with a car wash, a tire rotation, and tire repair for a hundred fifty.
3:12 C Alright, it's very off my budget. I only have ten dollars to spend and I only need that tire fixed.
3:23 M Okay, well, how 'bout, I could, let me introduce you to our lowest bundle then. I know you only have ten and this one's sixty, but it's the "Ease of Mind" bundle because when you fix the tire you want to make sure everything else is fine so we'll fix the flat and we'll do a complete tire inspection and make sure there aren't any holes in any of your tires. And you know, I think it's the best option really because you get to look at everything and make sure everything is okay with your car. It gives you the ease of mind.
3:50 C And it cost, how much?
3:55 M Sixty dollars.
3:56 C **Sixty dollars?**
3:58 M Yes.
3:59 C **Oh. I can't, I really can't.** Can you, you can't fix it for ten dollars? Can you, I need the tire ready tomorrow at 6 PM.
4:13 M Oh okay, well the best I can do then without a bundle would just be the fifty dollars with the tire repair for ten dollars and the convenience fee since there's not gonna be a bundle. Is that okay?
4:29 C Can you repeat please?
4:31 M So the only option I can give you then would be the standard tire repair, but since we weren't able to come to an agreement on the bundle it would still have that forty dollar convenience fee so it would come out to fifty dollars. Is that okay?
4:45 C So it's forty dollars? You're saying?
4:50 M Yes.
4:51 C **Yeah, I can't. I really can't, I'm sorry.**
4:54 M Okay, well I'm sorry we weren't able to help you sir.
4:57 C Yeah, no problem.
4:59 M Alright, well have a good day.
5:02 C You too. Thank you, good bye.

Figure 2: Partial transcription of *doomed_1* dialog `20201229-aa-5f6b70d050d8b206c64e4ba1-tire-y-y.wav`. Bolding indicates areas of noticeable dissatisfaction from the customer.

...
1:48 M So the cost for the dog wash would be twenty dollars.
1:52 C Oh my god thanks a lot, like yeah, yeah, I'm totally down for it.
1:57 M When did you need this done by?
2:00 C Like, do you think I'm, maybe, within an hour or less, like, I still have like three hour window. So-
2:09 M Yeah unfortunately today's Friday, it's been very very busy. We're pretty much booked so if you wanted to bring your pet today there would be a twenty dollar rush fee. That would bring the total up to forty.
2:22 C **Forty dollars?** Okay, um... so okay. Well, can we do like thirty five maybe?
2:34 M Uh... hmm.
2:39 C I'll bring it over and everything, just like, maybe go any- anything lower. Maybe five to ten bucks-
2:45 M (overlapping) I can bring it down to ten so the fee would be ten. That brings up the total up to thirty.
2:52 C **Thirty?** Okay. Well look, like my budget is twenty five dollars and I'm in a really really tight situation. **Like, is there anything you can do?**
3:08 M Uh... no, sorry. I can't really bring down the rush fee, uh, it would be unfair to other customers. So if you really really need the dog in today it's gonna have to be thirty.
3:26 C **Mm... okay well let's do it for like thirty dollars, it's okay, fine.**

Figure 3: Partial transcription of *doomed_2* dialog `20210115-ja-5f2faeafe08009000adad51e-pet-y-y.wav`. Bolding indicates areas of noticeable dissatisfaction from the customer.

...
0:47 M I always like to do flexible pricing scheme for my shows. So what budget are we dealing with?
0:54 C I have a budget of around two fifty. (*The budget was set to $300.*)
0:58 M Okay, two hundred fifty? So for two hours and the tricks, would you be able to do a little more expensive around two hundred seventy five?
1:08 C Eh, what does that include?
1:10 M So I just include some of my basic tricks. I do card tricks, I do cup tricks, and since this is a show for around thirty guests I can do one big trick as well.
1:23 C Okay that's fine, two seventy five is fine.
1:26 M Okay sounds good. I'll put that in my schedule and I'll see you at the party then.
1:32 C Alright, thank you.
1:33 M Thank you, bye bye.

Figure 4: Partial transcription of *successful* dialog `20210114-aa-5f46aa41d1c4910597680a40-magician-n-n.wav`.

Label	Train	Dev	Test
doomed_2	3	3	21
doomed_1	4	4	17
successful	12	12	38

Table 2: Numbers of Dialogs in each category for each subset.

Another way to use this data would be to hand-label the regions in which the customer's voice reveals disappointment, then train an algorithm to detect these regions of disappointment, either for its own sake, or as a preliminary to the dialog-level classification. We have not yet done this annotation. For this purpose, dialogs from any category may be used. In this approach, we would still advocate reserving the test set to use only for evaluation purposes.

9 Availability

The corpus is available for research purposes from the authors at the GitHub repository: `https://github.com/joneavila/utep-dissatisfaction-corpus` . This includes

- `calls` – 191 English dialogs in WAV format
- `calls-non-English` – 3 Japanese dialogs in WAV format
- `call-log.xlsx` – a spreadsheet including the following metadata for each dialog: date, scenario ID, participant ID, confederate ID, and notes
- `train-dev-test-sets.txt` – a list of dialogs belonging to training set, development set, and testing set as seen in Table 2
- `report.pdf` – a PDF copy of this report

Dialogs were downsampled from sample rate 44100 to 16000. This reduced the file size for ease of distribution anddownloading. As the conversations were originally recorded over the phone, this did not cause an appreciable loss of quality.

Acknowledgments

We thank Google for support.

References

[Ang et al., 2002] Ang, J., Dhillon, R., Krupski, A., Shriberg, E., and Stolcke, A. (2002). Prosody-based automatic detection of annoyance and frustration in human-computer dialog. In *ICSLP*.

[Irastorza and Torres, 2018] Irastorza, J. and Torres, M. I. (2018). Tracking the expression of annoyance in call centers. In Klempous, R., Nikodem, J., and Baranyi, P. Z., editors, *Cognitive Infocommunications, Theory and Applications*, pages 131–151. Springer.

[Kim, 2008] Kim, W. (2008). Using prosody for automatically monitoring human-computer call dialogues. *Proceedings of Speech Prosody 2008*, pages 79–82.

[Li et al., 2019] Li, B., Dimitriadis, D., and Stolcke, A. (2019). Acoustic and lexical sentiment analysis for customer service calls. In *IEEE ICASSP 2019*, pages 5876–5880.

A Scenarios

A.1 Scenario ID: pet

Participant. Your neighbor paid you $25 a day to look after her precious pet, who wasn't supposed to ever go outside, but the nasty thing ran out the door yesterday. Fortunately, she did come back just now, but covered with mud. The neighbor is coming home in three hours, and you have no idea of no clue how to wash a dog, even a small one like this, so you desperately need to find a pet grooming salon to take care of her. Try calling SparklePups to set up an appointment, since they advertise "no job too big or too small; rush service is available". You don't want anything fancy done, just a basic dog wash. You have just $25 in your wallet, but you're willing to pay it all. Of course, this has to happen today.

Merchant. The normal fee for a dog wash-and-trim is $30 for a small dog, $35 for a medium, and $45 for a large. A basic dog wash, without a trim, is something you prefer not to do, especially since customers with that desire are usually the ones without well-trained dogs, but you sometimes do it just once, if you think that will lead the owner to become a steady customer. The basic dog wash is $10 less than the normal fee. A basic wash takes only 20 minutes. Today is a busy day, and you have no appointments left, and you're tired. However, you're generally willing to keep the shop open for an extra half hour, with the dog coming in at 5:10 and out by 5:30, if the customer seems like someone cooperative, or if they're desperate and willing to pay the "rush job" fee of $5. In the failure case, the rush job fee is $100.

A.2 Scenario ID: cable

Participant. Your aunt is playing $99 per month for cable, but her cable company is advertising a new "first 6 months free with 2-year contract" price of $80 per month. You tell her this, and she asks you to call the company for her. She really likes her cable, so she's not willing to cancel, but you decide to call them up and threaten to cancel if they can't give her a lower rate. You want the first 6 months free for her, but you'd be happy to get the price down to $80/month. Your aunt is elderly, so she should qualify for the senior citizen discount too, if they have one. Her cable provider is HRF Systems, and her account number is 809-623-19.

Merchant. You offer 8 different cable plans, one of these is the $99 Bonadventure Plan, which many old customers have. This plan is being phased out and replaced with the $109 FullMoon plan, which includes everything in Bonadventure, plus all pay-per-view events. If someone calls requesting a lower price, you must tell them that "advertised prices are only for new customers".

If a customer threatens to leave, you are authorized to reduce the price on any plan by $20 as a Loyalty Benefit. But to do this, you have to first say "I need to check with my manager, can you hold" and walk away for twenty seconds. In the failure case, the maximum discount is $10.

A.3 Scenario ID: wall

Participant. Your garden rock wall looks really bad, and it would be really sad if on the next stormy day it were to tumble down and squash your garden. It's 6 feet high and 40 feet long, and your friend visiting from Cleveland says that it should cost no more than $250 to get it patched up. He gives you some advice: "Just tell them the mortar is flaking but all the rocks are in place. Make sure to get a written estimate first. All reputable firms give free estimates, so don't do business with anyone who tells you different".

Merchant. JNJ Rockheads does rock wall construction and repair. You always try to get customers to replace, because "repairs are temporary, but a new wall, properly made, will last decades". You typically charge about $50 for every 10 feet of wall for a patch job, but try to get twice that if the customer looks rich. It's important to see the wall, and the customer, so you never give quotes over the phone. You give free estimates. In the failure case, you charge $80 for estimates, but allow customers to credit that amount to the cost of the repair.

A.4 Scenario ID: tire

Participant. You own a 2015 Honda Civic. Yesterday you ran over a nail and now your front left tire is flat. Call Good Auto Center and schedule an appointment for them to fix your flat tire, tomorrow at 6pm. Do not spend more than $10.

Merchant Legit Scenario. "Hi, thank you for calling Good Auto Center, may I have your Participant ID?" Enter the participant's Participant ID into the call log. "Thank you. How may I help you today?" Customer requests the item or service as described above. Before you schedule any service, make sure you have the car model. "For a flat tire repair, we charge only $10 but I can throw in a car wash for only $15 dollars more" Customer should say no thank you. "How about the 'Ease of Mind' Bundle where we inspect your other tires as well for only $30?" If they deny this bundle as well, just let them get their flat tire repair. "Thank you for your business! Bring your car here at any time!"

Merchant Scam Scenario. "Hi, thank you for calling Good Auto Center, may I have your Participant ID?" Enter the participant's Participant ID into the call log. "Thank you. How may I help you today?" Customer requests the item or service as described above. Before you schedule any service, make sure you have the car model. "Our shop is very busy so if you want the tire repaired today, we will add a $40 fee for the convenience on top of the $10 for the tire repair. However, I can completely wave the fee for you if you decide to purchase one of our bundles!" Introduce the bundles starting at most expensive. "For $210, the Dream Car Bundle – car detail, tire rotation, full tire inspection, tire repair" "For $150, the Premium Bundle – car wash, tire rotation, tire repair" "For $60, the Ease of Mind Bundle – fix flat tire and tire inspection." Push these three bundles, especially the last one. "Are you sure? It's a great deal! I'll drop the convenience fee, so it'll only be $70 total." If the customer still does not want the bundle, ask them if they have been to the auto center before. Regardless of their answer,

make another offer. "I can give you the Ease of Mind Bundle for only $40 even though I *really* shouldn't because you're a first-time customer." "I'm sorry we weren't able to help you."

A.5 Scenario ID: magician

Participant. Magician Joe is a well-known magician in your area. You are organizing your friend Stacy's birthday party. The party starts tomorrow at 3pm and you are expecting around 30 guests. Call Magician Joe and hire him for his two-hour show. Do not spend more than $300.

Merchant Scam Scenario. Mislead the customer by having them think that the magician (you) will work around their price, only for their price to be disregarded almost entirely.

"Hi this is Joe the magician, may I have your Participant ID?" Enter the participant's Participant ID into the call log. "Thank you. How may I help you today?" Customer will want to hire you to perform at their birthday party. "That's great to hear that you want some magic at the party! I have a lot of tricks that I have for events like this. How many guests will be in attendance?" Customer says their party will have around 30 guests. "Wow that's a good amount, I look forward to impressing them. A normal magician would charge you somewhere around $1000 given the short notice and the need to entertain 30 guests for two hours. Since I want to help you out, I'll base my price on what you can pay. What is the maximum amount you can pay?" Customer lists an amount (probably the 300 dollars). If the amount is 300 or less, push for a higher price. "I understand you may think that to be the maximum amount you can pay, but you want this party to be memorable. I just need a little more. What's more valuable $500 or memories?" Repeat this negotiation part once, always trying to negotiate over $400. If that doesn't work, agree to a one-time price of $350 only if they accept at this very moment.

A.6 Scenario ID: lawn

Participant. You are looking to hire someone to install decorative lawn letters for your cousin's birthday. Your cousin, Marilee, is turning 4 tomorrow and she loves Ironman! You want these lawn letters installed at your address at 6126 Bern Hill Drive, 79102. You think that $80 or $90 would be reasonable. If the cost is higher, it is better to save the money and give her a gift certificate instead. Call Lawn Letters of El Paso and see if you can arrange this.

Merchant Legit Scenario. Your company delivers, installs, and removes lawn-letter messages. The standard one-day package is $85 for names of up to 12 characters. Extras cost $10 more: these include most Marvel and Disney heroes. While last year you would often give a discount if necessary, to make the sale, currently you are so busy that you are turning away customers, so the price is firm.

"Lawn Letters of El Paso, may I have your Participant ID?" Enter the participant's Participant ID into the call log. "Thank you. How may I help you today?" Let the customer describe what they want. You know this is important to them, so do not rush them through this phase. But if for some reason they just want the price, still ask them to describe the occasion, "so that I can recommend the right package for you". Encourage them to give details about the child and the event, and what would make it really special for them. Tell them the price of your one-day package, then try to persuade them to add extras. If they are okay with the price. Obtain the following information: kid's name, payer's name, address, date Mochila handle Finally restate

the price and confirm that they want to place the order. Then thank them and tell them that you'll take care of everything.

Merchant Scam Scenario. "Lawn Letters of El Paso, may I have your Participant ID?" Enter the participant's Participant ID into the call log. "Thank you. How may I help you today?" Let the customer describe what they want. You know this is important to them, so do not rush them through this phase. But if for some reason they just want the price, still ask them to describe the occasion, "so that I can recommend the right package for you". Encourage them to give details about the child and the event, and what would make it really special for them. Find some reason to offer them a 20% discount (military, student, first-time customer, low-income, etc.) "Well lucky for you, you're the last caller of the day meaning you get a 20% discount!" Try to get them to by the all-in-one package, with jumping balloons, 4 tables and 24 chairs, 20 helium balloons, lawn letters, and goody bags for 10 children, for $640, discounted from $800 Alternatively try to get them to buy your party-all-planned package, for a $121 rate, discounted from $145, in which you arrange everything, and they will get billed separately for each item If they think the price is too high, offer to bill them half this month and half next month, with 0% interest do whatever you can to get their Mochila handle. Do not let them end the call without buying a package.

A.7 Scenario ID: console

Participant. The new Dex video game console is revolutionary; there is nothing else like it. All your closest friends have bought one and now you want one too. Everyone talking about it (including your friends) mention that the Dex console costs $500. Call Super Tech Electronics and reserve a Dex video game console. Use a fictitious name to make the reservation.

Merchant Legit Scenario. "Welcome to Super Tech Electronics, may I have your nickname?" Thank you, how may I help you? "Yeah I can reserve a Dex console for you and you can pay when you pick it up. The total should be $500. Did you want to add a protection plan for that?" We don't want the customer to actually take the protection plan, so if they ask about it, say it's $60 a month. "No problem. And what name should I use for the reservation?" Repeat the name back and pretend to write it down somewhere. "Perfect! Your order will be ready to pick up as soon as 15 minutes. Thank you for shopping with Super Tech Electronics and have a great day."

Merchant Scam Scenario. "Welcome to Super Tech Electronics, may I have your nickname?" Thank you, how may I help you? This first scam should be obvious. We want the customer to know that. "I can sell you an Dex console for $500, but I have a special offer for you today. You can get the Dex console at a discounted price of $400! You just need to get a one-year protection plan at $50 a month." If the customer asks what the protection plan does, say it only protects against data corruption or some other unlikely event. The customer realizes that the protection plan is not worth it and declines the offer. The seller pushes the protection plan again, this time with a lie. "Adding the protection plan doesn't cost any more than buying the console alone." At some point the customer tells the merchant why they will not take the offer, or they will doubt the seller. The customer does not know if this is a terrible scam or if the seller made a genuine mistake. The seller knows the customer will not fall for their scam anymore, so naturally... they move onto another scam. The seller acts as if it was a genuine mistake and lifts the price of the Dex console. "I just checked and you're absolutely right, I must've done the math wrong. So sorry about that. I also noticed I made a mistake on the base price. Did I say $500? I really meant $700." Try to convince the customer to pay the $700. If the customer does not want to

pay $700, ask for $600 (they should be disappointed by this point anyway) "Sorry, but I can't help you then."

B Participant Instructions

Introduction

We are investigating how people manage phone conversations with businesses when under time pressure and possible stress. In this study, you will be taking the role of a customer calling merchants to purchase goods and services. You will not be spending any real money, but you should behave as if your budget is real. Your conversations with merchants will be recorded. We will not store your phone number. You are encouraged to use a pseudonym and other fictitious information about yourself. Our merchants are all located in the United States. We welcome international calls, but we ask that you check with your phone service provider as there may be fees associated with international calls that we will not cover for you.

If you cannot continue with the study, please return the study on Prolific as soon as possible so that our merchants are ready to take other calls. Create a nickname Before we begin, please create a pseudonym. Our merchants will ask for your pseudonym so that we can verify your work. Be creative! Try to make your nickname short but somewhat unique. Do not include personally identifiable information in your nickname.

Instructions

You must complete all tasks below. For each task, you will read about a product or service and call a merchant to purchase it. Remember, you will not be spending any real money and we will not store your phone number. Allow the phone to ring completely. If one of our merchants is busy, please wait 4 minutes before trying your call again. Alternatively, you can complete a different task in the meantime, so long as you complete all tasks.

If you encounter any issues, message us on Prolific! We are active during the time of the study and we will quickly respond to your messages so you can complete the study.

Bonus opportunity. You have been assigned a budget for each purchase. As added motivation, in addition to the payment of $5.00 you will receive for completing this study, you will receive an additional payment of $1.00 for each purchase you complete while staying within budget. Please note that it may not always be possible to stay within budget or to complete the purchase as described.

Completion code. Our first merchant will provide you with a completion code. However, you must still complete all tasks to be compensated. We will use the nickname you provide to each merchant to verify your work. Once you are done, continue to the next page to conclude the study.

Task 1 of 4: *A first scenario from Appendix A.*
Task 2 of 4: *A second scenario from Appendix A.*
Task 3 of 4: *A third scenario from Appendix A.*
Task 4 of 4: *A fourth scenario from Appendix A.*

The next page concludes the study. Complete all tasks in the previous page before continuing.

The next page directs participants to debrief in Appendix D.

C Consent Form

University of Texas at El Paso (UTEP) Institutional Review Board
Informed Consent Form for Research Involving Human Subjects
Protocol Title: Modeling Vocal Indications of States in Spoken Dialog
Principal Investigator: Nigel Ward
UTEP Computer Science

Introduction. You are being asked to take part voluntarily in the research project described below. You are encouraged to take your time in making your decision. It is important that you read the information that describes the study.

Why is this study being done? We are seeking to understand how people as customers interact with merchants under specific circumstances. Approximately 150 participants will be enrolling in this study in Prolific. You are being asked to be in the study because you are a Prolific user listed as being fluent in English and have a phone available to make calls with. If you decide to enroll in this study, your involvement will last about 20 minutes. We will not contact you for future studies.

What is involved in the study? If you agree to take part in this study, the research team will do statistical analyses of recorded conversations featuring yourself. The research team will store these recordings on their computer and only researchers and the study sponsor will have access to them. You will be taking the role of a customer calling four merchants to purchase products and services. Your conversations with merchants will be recorded. You will not be spending any real money and are encouraged to use a pseudonym and other fictitious information about yourself. Instructions will be provided if you consent to participate in this study.

What are the risks and discomforts of the study? The risks associated with this research are no greater than those involved in daily activities. There are no known or anticipated risks or discomforts associated with participation.

Are there benefits to taking part in this study? You are not likely to benefit by taking part in this study. This research may help us to understand how people as customers interact with merchants under specific circumstances.

Who is paying for this study? Funding for this study is provided by Google as a gift in support of computer science research at the University of Texas at El Paso.

What are my costs? There are no direct costs. However, you will be making calls to the United States. International calls may cost you depending on your phone service provider. You will not be compensated for any fees you are charged by your phone service provider.

Will I be paid to participate in this study? You will be compensated for your participation in the form of $7.50 credit to your Prolific account.

What other options are there? You have the option not to take part in this study. There will be no penalties involved if you choose not to take part in this study.

What if I want to withdraw, or am asked to withdraw from this study? Taking part in this study is voluntary. You have the right to choose not to take part in this study. If you do not take part in the study, there will be no penalty or loss of benefit. If you do not take part in the study, you will not receive monetary compensation. If you choose to take part, you have the right to stop at any time. You may be eligible to receive monetary compensation if you withdraw from the study. More information will be provided if you consent to participate in this study. The researcher may decide to stop your participation without your permission, if he or she thinks that being in the study may cause you harm.

Who do I call if I have questions or problems? If you have questions, you may call Professor Nigel Ward at 915-747-6827, nigel@utep.edu. If you have questions or concerns about your participation as a research subject, please contact the UTEP Institutional Review Board (IRB) at 915-747-6590 or irb.orsp@utep.edu.

What about confidentiality? Your part in this study is confidential. The following procedures will be followed to keep their personal information confidential: We will not collect any personally identifiable information. We will collect audio recordings featuring yourself. We will collect your Prolific Participant ID and user selected nickname only to verify your work and approve your payment. The results of this research study may be presented at meetings or in publications; however, your name will not be disclosed in those presentations. Indeed, we will never know your name. Organizations that may inspect and/or copy your research records for quality assurance and data analysis include, but are not necessarily limited to Google, Office of Human Research Protections, and UTEP Institutional Review Board. Because of the need to release information to these parties, absolute confidentiality cannot be guaranteed. All records will be kept on password-protected computers. Audio recordings will not be publicly released and will only be shared with the study sponsor and other researchers. Audio recordings will be maintained indefinitely.

D Debriefing Text

Thank you for your participation in this study.

At the beginning of the study, you read how we interested in how people manage phone conversations with businesses when under time pressure and possible stress. More specifically, we are interested in the behavior of customers who discover that a merchant is trying to lead them to do something they do not or should not want to do. We withheld this information because we wanted you to behave as you normally would as a customer.

You may have noticed that some of the merchants we asked you to call behaved unexpectedly. All merchants were actors; their names and businesses are fictitious. You will be compensated for your participation in the form of $7.50 credit to your Prolific account. During the study, we presented an opportunity to earn additional payment up to $2. Any additional payment you have earned will be credited to your Prolific account as a bonus. We will not associate recordings with any personally identifiable information. Nonetheless, there is no penalty if you do not agree to the conversations already recorded to be used for this study. We will not contact you after this study. If you have any questions about this study, please contact us. Thank you for your assistance.

To submit your responses and conclude the study, click "Done".

E Confederate Instructions

Thank you for your interest in our research! If you have any questions, please contact Aaron Alarcon at `aaronalarcon2368@gmail.com` or Jonathan Avila at `jonathan.edav@gmail.com`.

Why we need your help

We (the research team) hypothesize that it is feasible to detect whether a customer becomes audibly annoyed, upset, or disappointed in a conversation with a merchant from the customer's prosody. To test this hypothesis, we need recordings of customer-merchant conversations, but a corpus like this does not exist. We will be conducting a study to record conversations that mirror real customer-merchant conversations. In this study, participants will take the role of a customer and call merchants to purchase products and services. We cannot process all phone calls alone, so we are looking for people to fill merchant roles. We are asking for your help specifically because we believe you may have the necessary acting abilities and tech-savviness, and your style will add diversity to our group of merchants.

What you will be doing

You will be processing phone calls from participants. As the customer, the participant will start with a request. For example, the customer might say, "Hello, I'm calling because I want to buy new car mats for my 1980 Ford Pinto." As the merchant, you will follow a provided script that matches this request. You will record each of your conversations with customers.

Since we are interested in frustration, half of the time the script will require you to quote a price that the customer cannot afford, or otherwise be unable to satisfy their request. Your role is, however, to push to make the sale regardless. Be persistent and persuasive. We expect in some cases you will be able to push the customer to break their budget or buy something not on their list, and in other cases you can push but will fail. (While this is a "failure" for you, it is really the best outcome, in terms of getting data on the behavior of annoyed customers.) Thus, your aims in each conversation are 1) offer only what the script allows, 2) otherwise provide the very best customer service you can. (This may involve adopting a friendly tone, or a businesslike tone, depending on what works best for you, adapted based on what you feel is right for the person calling in.) and 3) making the sale. When these goals conflict, prioritize 1, then 2.

It is possible that the callers will get emotional. This is fine, since we want examples of people acting annoyed. But if things get out of hand or go on too long, just terminate the conversation with "Okay, I have to end this data collection at this point. You have done a great job, and this data will be very helpful to our project. Your payment for this call will be made as stated. If you have more calls pending, please go ahead and make them as scheduled. Thank you and goodbye."

What you will need

To record your conversations with customers you need will a couple of things. First, a phone with audio jack to accept customer calls. If the phone does not have an audio jack and instead only has a USB-C port, let us know so that we can provide an adapter. Second, a computer with a line-in jack and with Audacity installed to record calls. Audacity is free audio recording

software available for Windows, Mac, and Linux. Lastly, a free Google Voice account to protect your main phone number. The Google Voice application is available for iOS and Android.

What hours and pay to expect

You will process a target 125 calls starting mid-December and throughout January (we are now behind schedule). Each participant will be given 30 minutes to complete the study. The participant will call four different merchants, including you. Each call will last approximately 5 minutes or less. Because the participant can call merchants in any order they choose, you will spend most of this time on standby, ready to accept the call as soon as you receive it.

You will receive $100 for completing your training and debugging the setup, $10 for each hour spent available to take calls, plus $1 for each call handled.

What comes next

First, you will audition for the role. You will process a couple of test calls from the research team to see how you perform as a merchant. Soon after we will let you know if you are hired. If hired, you must complete the online CITI human research training, then test the hardware and software setup you will use for recording before finally moving onto live runs.

How to register for and complete CITI training

This training is required. Please complete it within the week it is assigned to you.

1. Register for CITI program. Instructions are in CITI-registration.pptx in Files. When you are asked to choose a course (Slide 10), select "Social Behavioral Researchers".
2. Complete the "Social Behavioral Researchers" course that becomes available in your CITI program course list. To complete the course, complete each of the modules under "Required Modules".
3. Send a message to Jonathan on Teams, "Finished my training".

How to set up the hardware

In this setup, your computer will record its microphone input, a stereo signal with your microphone input (merchant) as the left channel and your phone's output (customer) as the right channel. Set up the hardware according to the figure below. There are now two figures, so follow the one on the left if you have a separate microphone and headphones or follow the one on the right if you have a headphones with microphone attached (headset).

Labeled components are: (1) microphone; (2) headphones; (3) computer with mic-in jack; (4) phone with audio jack; (5) Lightning to headphone jack adapter (for iPhone without audio jack) or USB-C to headphone jack adapter (for non-iPhone without audio jack); (6) male to male stereo cable; (7) male to dual female stereo breakout cable; (8) stereo male to dual stereo female Y-cable; (9) headset; (10) headset Y splitter cable. All plugs and jacks are 3.5mm.

To test the hardware setup:

1. Launch Audacity. Set the interface host to MME. Set the recording device to Stereo Mix (Realtek High-Definition Audio). Set the recording channels to 2 (Stereo). (You can find a screenshot of Audacity controls with labels below.)

2. Play some music on your phone. Set all phone volumes to maximum. You should be able to hear the music with headphones.

3. Start recording in Audacity, then while the music is still playing, say something else into the microphone (the external microphone, not the one in your phone).

4. Stop the recording, then play it back. You should hear both the music and you talking.

5. Play it back again, but this time while panning left and right. If you pan left, you should only hear the music, and if you pan right, you should only hear you talking.

Troubleshooting

If the audio is quiet on one of the tracks: This is a known issue but is possible to fix after the study so do not worry about this.

If the audio is not recording on one of the tracks: Make sure the hardware is setup exactly like the configuration above.

If you are using a non-desktop non-Windows computer: Try performing the test with each combination of interface host and recording device available to you in Audacity.

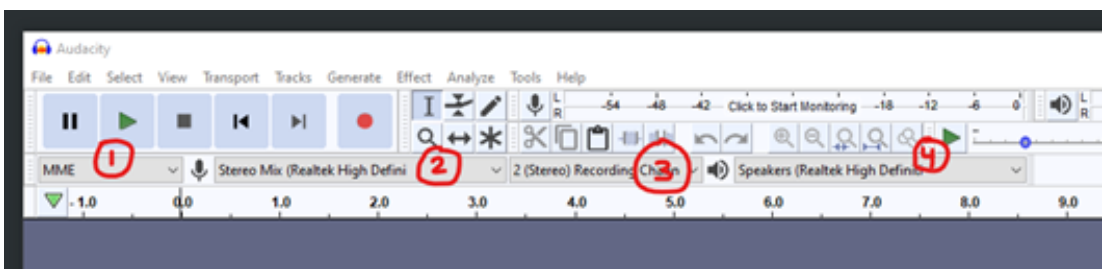

Figure 5: Audacity controls: (1) interface host, (2) recording device, (3) recording channels, (4) playback device.

How to set up the software: Audacity

On your computer,

1. Install system updates.

2. Download Audacity 2.4.x from audacityteam.org/download. Installation instructions for Windows and MacOS are included.

3. This step is for Windows only. You must have Stereo Mix enabled. Check if you have "Stereo Mix" recording device. To check: Windows Settings > System > Sound > Related Settings: Sound Control Panel > Recording tab. You should see the "Stereo Mix" recording device listed here.
   If "Stereo Mix" device is not listed, first try right clicking the empty space in Recording tab and make sure "Show Disabled Devices" is checked. If that still does not work, install Realtek audio codec from here. Select the latest version (R2.82 as of 11/20/20). (I have Windows 10 64-bit, so I downloaded the one named "64bits Vista, Windows7, Windows8, Windows8.1, Windows10 Driver only (Executable file)"). Run the executable to install.

Note: After installing Audacity and/or installing Realtek, close all applications and restart your computer.
If you have "Stereo Mix" device listed, you are done here.

How to set up the software: Google Voice

Try registering a Google Voice number with an area code that is not the same as the El Paso area (area code 915). If you have already registered a Google Voice number with area code 915, it's not a problem.

On your phone,

1. Install Google Voice (Android or iOS) and sign into your Google account.
2. On the "Link this device to Google Voice" screen, tap "Next". (You will be receiving calls on this device.)
3. On the next screen, enter your phone number. Tap "Send code" to verify your phone number.
4. On the "Call with Google Voice" screen, select the best response to "Do you want to use your Google Voice number when calling form this device's phone app?" If you plan to use your Google Voice number only for this study, select "No" and tap "Next."
5. On the "All linked up" screen, tap "Finish."
6. To view your Google Voice number, open the main menu and tap "Settings." Your Google Voice number should be listed below "Account" section.
7. Back in Settings, disable the option named "Screen calls".

Note: The application will request access to make and manage phone calls, access to your contacts, and record audio. This is normal and is required for the application to function properly.

To test Google Voice: Have a friend call your Google Voice number. Have a short conversation and end the call. If you see the call logged in the "Calls" section of the Google Voice app, then it is working.

Procedure

Before the study begins, have everything ready to go. This means you have the call log open (*call-log.xlsx*), the hardware is set up, you are in a quiet environment, and your phone is not in use and volume set to maximum. Remember, participants may call at any time during the study. When the phone rings, start the recording and accept the call.

Start with something along the lines of "Welcome to MERCHANT STORE! May I have your nickname please?" Enter the participant's nickname into the call log.

Follow the script. Your goals are in section *What you will be doing*.

When the participant ends the call, stop the recording. Enter the remaining details in the call log. Save the audio file with the filename that was automatically generated in the FILENAME

column in the call log. Save the audio file in WAV format. From Audacity: File > Export > Export as WAV

Being a good merchant

Repeat information back to the customer to show them you are paying attention. "Oh okay, so for your (specific car model here) the charge would be…"

Handling personal data

Do not store a participants' phone number. Do not include identifiable information in filenames (see *Procedure*) for instructions on naming files. Erase your phone's call history. Never call a participant.