

Designing a GWAP for Collecting Naturally Produced Dialogues for Low Resourced Languages

Zulipiye Yusupujiang, Jonathan Ginzburg

Université de Paris, CNRS, Laboratoire de Linguistique Formelle
zulipiye.yusupujiang@linguist.univ-paris-diderot.fr, yonatan.ginzburg@univ-paris-diderot.fr

Abstract

In this paper we present a new method for collecting naturally generated dialogue data for a low resourced language, (specifically here—Uyghur). We plan to build a games with a purpose (GWAPs) to encourage native speakers to actively contribute dialogue data to our research project. Since we aim to characterize the response space of queries in Uyghur, we design various scenarios for conversations that yield to questions being posed and responded to. We will implement the GWAP with the *RPG Maker MV* Game Engine, and will integrate the chatroom system of the game with the Dialogue Experimental Toolkit (DiET). DiET will help us improve the data collection process, and most importantly, make us have some control over the interactions among the participants.

Keywords: GWAPs, Response Space, DiET, Low Resource Languages

1. Introduction

There are many tasks such as data and text annotation, and image labeling, which challenge even the most high-performing computational algorithms or computer programs, but which are actually very easy to handle for humans. This motivated (Von Ahn and Dabbish, 2008) to propose a method called “games with a purpose” (GWAPs), in which players produce a large amount of useful data and perform tasks which are difficult for computer systems while they are entertaining with interesting online games. As a result, players make considerable contributions towards providing a sufficient amount of annotated data for training and developing computational algorithms. GWAPs have been used successfully in many tasks such as *the ESP Game* for image labeling (Von Ahn and Dabbish, 2004), *Peeka-boom* (Von Ahn et al., 2006) for locating objects within images, and also the *FoldIt* (Cooper et al., 2010; Khatib et al., 2011) for protein folding. GWAPs have also been used for solving NLP problems, for instance the *Phrase Detectives* (Poesio et al., 2013) for creating anaphorically annotated resources, and also the *Puzzle Racer* and the *Ka-Boom!* GWAPs for word sense disambiguation.

The main objective of our project is to characterize the response space for queries across languages. A question can be responded to in many ways. (Łupkowski and Ginzburg, 2016) studied one significant component of the response space of questions, which is responding to a question with a question. They used the British National Corpus (BNC corpus) and three other more genre-specific corpora in English, and classified the range of question responses using 7 classes (Clarification requests, dependent questions, questions about the form of the answer, requests for underlying motivation, indirect question responses, and two classes of evasion questions), and showed how to model these 7 classes within the framework of Conversation Oriented Semantics (KoS in short), which is based on the formalism of Type Theory with Records (TTR) (Ginzburg, 2012). Furthermore, (Ginzburg et al., 2019) offered a characterization of the entire response space for English and Polish. Based

on these works, we aim to take the challenge of characterizing the response space of questions in low resource languages that have yet to be studied in this regard. However, most low resourced languages have very little or even no digital language resources available for conducting scientific research. In particular, a sufficient amount of naturally produced dialogue data is rare to find. Therefore, constructing a dialogue corpus for such low resourced languages is the first essential step of our project. In this project we focus on collecting dialogue data and analyzing the response space of questions in Uyghur since it is a low resourced language under threat. In addition, conducting field work for collecting Uyghur dialogue data is not possible given the current difficult situation of the target area. Thus, we propose a new method for collecting natural dialogue data using a GWAP.

Inspired by current work on GWAPs, we plan to design and create a Massively multiplayer online role-playing game (MMORPG), using the **RPG Maker MV Game Engine**.

¹ Since MMORPG involves a large number of online players and their active interactions with each others, through text or voice, we hope to collect the naturally produced dialogues that take place during each game.

In this paper, we aim to present our initial design and further plan for this GWAP for collecting dialogue data. In the following section, we introduce the initial design for how to get players to produce a large variety of natural dialogues while playing the game; we then introduce the chatroom system of our GWAP and the data collection process; subsequently, we give the overall rules and conventions of the game, and also the reward system; finally, we draw some brief conclusions and offer some ideas for future work.

2. Integrating Task-oriented Dialogue with GWAP

Since our project is about characterizing the response space of questions, we hope to collect a varied corpus of questions and answers. Therefore, we have to pay much attention to enticing the participants to produce different

¹https://store.steampowered.com/app/363890/RPG_Maker_MV/

types of questions and answers during the game. Consequently, our primary goal is to ensure that the players chat extensively on a wide range of topics. To achieve this, we plan to design some task oriented dialogues inside the GWAP so that players will participate in dialogues with more specific guidelines and certain aims. The spoken part of The British National Corpus (BNC), which is used in (Łupkowski and Ginzburg, 2016; Ginzburg et al., 2019), was collected in different contexts, including formal business meetings, government meetings and also radio shows and phone-interviews. To evaluate the difference between the data collected by our method and the BNC corpus, we will also use our GWAP for collecting English dialogue data so that we can conduct a comparative study on two different English dialogue corpora.

To achieve our goal, we plan to design and implement an MMORPG using the RPG Maker MV Game Engine. We will create a virtual world in our GWAP, so every player of the game has several options for choosing a character class and role-playing class in the virtual world. Virtual worlds (Bartle, 2004) are computer systems or applications which imitate a real environment. They can be simultaneously affected by an enormous numbers of players, and can exist and develop internally. Thus, virtual worlds are said to be *shared* or *multi-user*, and also *persistent* (Bartle, 2004).

The game *Ring Fit Adventure*,² is a game in which players accomplish their fitness exercises while being entertained with an adventurous and fun game. It is a Role Playing Game in which there is a big world to explore, monsters to beat, and also bonus items to collect. Players can freely choose the level of exercises and manage to accomplish different work-out exercises during their journey in the virtual world. Since the main objective of this game is to encourage or force the players to do more exercises, it is designed in a way that players have to do different exercises in order to battle various monsters they come across on their way, otherwise they cannot continue their journey. This game has become very popular since it is a very new and fun way to keep one motivated to reach a fitness goal.

Inspired by the *Ring Fit Adventure*, we have come up with a similar idea which encourages or in a sense "forces" the players to have a discussion on various topics with other players during the game. Players will be given several topics to choose, or sometimes will be randomly assigned to a specific topic, and they will have a free chat according to the instruction within a time limit. In what follows, we sketch the initial design of scenarios for various tasks in the game:

- **Role-playing:** in this task, you will be role playing one of the characters in the following story: There was a severe public attack yesterday on the main street of your city. The police has successfully arrested one of the assailants, who is currently being interrogated.
 - If you are the police: you should ask various questions from the suspect to obtain a confession,

and try to let him admit his crime, and also force him to disclose his accomplices;

- If you are the criminal suspect: you should try your best to deny your crime, and make the police believe that you are innocent.
- **Planning a given task:** you are invited to participate in a real-life TV show, and you are paired up with a stranger (who is also here to participate in the TV show). Your task is to plan a trip together to a totally unfamiliar place. The two of you should work together on planning the entire two week trip. The trip is self-funded so you may want to discuss your financial situation and how to arrange the budget for the trip. Since you will travelling together for the entire two weeks you should start by getting to know your partner well, including his/her basic information, family situation, hobbies etc.
 - **Direction giving:** in this task you will be chatting with your partner in order to find out how to get to his/her current address. You should pay close attention to the details and draw a travel plan to your partners's place. This task should be done in two rounds, with each of you playing both roles.
 - **Real open discussions:** in this task, you and your partner/partners in the chatroom should freely and openly discuss a topic, you may discuss an issue happening around the world, or a news item, politics, comedies, education, or anything you may interested in. During the discussion, you should ask each other various questions about the topic.
 - **Future ideal society:** in this task you and your partner/partners in the chatroom should discuss the ideal future society you want to live in. You should tell your partner how is your ideal future society, and your partner should ask questions about that ideal society. You can talk about the social system, education, medical, transportation, and any other aspect of that ideal society.
 - **Interviewing:** in this task, you will be role-playing an interviewer or an interviewee.
 - If you are the interviewer, you should ask various questions of the person you are interviewing, including basic information, private information, their current mission, their opinion about some topics, or even their further plans.
 - If you are the interviewee (you will role-play one of the famous people randomly assigned to you from our list), you may choose quite freely how to respond, you may want to answer correctly, or lie to the interviewer, you can refuse to answer or change the topic.

²<https://www.nintendo.com/games/detail/ring-fit-adventure-switch/>

- **Guessing other person’s current location:** in this task, you will guess the current location of your partner according to the description of their surrounding environment. You can also ask some questions to verify your partner’s location, such as, ‘*Is there a desk nearby?*’. You should ask as many questions as you need to correctly guess the current location of your partner. For the safety reason, the current location does not have to be the players’ real location so they can make up or just imagine a place and let the others to guess.

3. Design of the Chatroom in the Game

All conversational tasks above will take place in the chatroom system which is implemented in the game. Each time the players are assigned to perform a task, they will be able to choose chat-rooms with different numbers of participants. We will have chat-rooms with only two players, with 4 5 players, and also with 6 10 players. This will enable us to make comparisons of dialogical behaviour across 2 person dialogue, small group multi-party, and larger group (Carletta et al., 2000; Ginzburg and Fernández, 2005). Players are randomly assigned to the chat-rooms to prevent cheating, and they should actively participate in the discussion until the time expires. To guarantee a successful dialogue flow, players will be informed at the beginning about the conversation rules and typing conventions, such as adding a question mark after each question and the minimal number of turns they should take during the conversation. In the end, after using a text mining technique, an overall assessment of the conversation will be sent to all participants of the conversation. They will be rewarded if they fulfill the requirement, otherwise they have to participate in another conversation. Since we are interested in various types of questions and answers with different word using, sentence making and speaking style, players will have freedom of speech so they are able to discuss any social issue or topic without restriction. This can be rather intriguing for people who used to live in a society where the government strictly controls their speech and actions. The real-world private information of the players will not be provided during the game, therefore, players should feel safe and free to discuss.

Furthermore, in order to collect conversational data in a more efficient way, we will adapt the Dialogue Experimental Toolkit (DiET)³ (Healey et al., 2003) to be usable on the internet and link it with the chatroom system of our game. DiET is in its original form a text-based chat-tool which allows utterances of particular types to be artificially introduced into natural dialogues in a controlled and synchronous manner, and unknown to the dialogue participants. DiET lead to novel findings about dialogue interaction (Healey et al., 2003; Eshghi and Healey, 2016). This methodology has been approved by several ethics committees (e.g., at Stanford), as long as the subjects are debriefed after participation, similarly for online GWAPs⁴.

³<https://dialoguetoolkit.github.io/chattool/>

⁴<https://www.scienceathome.org/legal/game-consent-form/skill-lab-science-detective-terms-of-service/>

WORKFLOW OF DIET

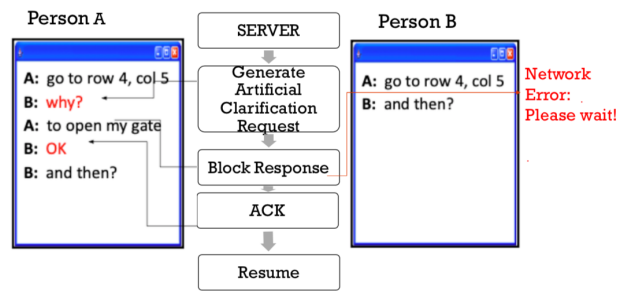


Figure 1: Example of the workflow of DiET chattool

The DiET chattool can automatically record all activities of the participants including key-presses, words and turns, typing notifications, read receipts, number of edits, typing speed, typing overlap. All dialogue data is immediately saved in various formats so it can be loaded into different data analysis tools such as Excel, SPSS, R, MATLAB, etc. Therefore, there is no need for transcription and post-processing of the data. Integrating the chatroom system of our GWAP with DiET helps us improve the data collection accuracy and efficiency.

Another key strength of DiET is that it can manipulate the interactions between participants. As seen in 1, DiET can send artificial clarification requests such as “why?”, “what?”, some fake feedback “Ok”, “yeah”, and also some artificial hesitations such as “umm”, “uhh”. These interactions are unknown to the participants so it can assist participants in producing more dialogues and various types of question responses which we are interested in. Thus, DiET’s integration with our chattool system can help us reduce the problem of data sparseness.

4. Gameplay and Mechanics

Since the Uyghur diaspora outside China who are above 16 years old are the primary target audience of this game, we would like to attract more Uyghurs to participate through designing the game to their taste. Therefore, we will use some famous Uyghur fairy tales as narratives of the game and also use the original names in these stories as some characters’ names in the game. Since most of the Uyghurs are familiar with those fairy tales, it will hopefully be intriguing for them to play the game. In addition, we will develop the game in English and in Uyghur, so Uyghur can also be the operating language of the game. To the best of our knowledge, there is almost no game which has Uyghur as the operating language. Having a game in their mother tongue would be novel and interesting for the target audience, so there will be more people attracted, especially people who are not familiar with the main world languages, and will give the game a try. Most importantly, the Uyghur diaspora are well aware that their language is under threat and they have to make a great effort to keep it alive. Thus, participating in such a game with a scientific purpose will potentially impress members of this community.

We will build a MMORPG with a linear story structure and the game moves on as the improvement of player’s lev-

els and skills. There will be several game sections separated by the difficulty levels, so the players should gain enough experience points and skills to move to the next higher game section. To begin with, players should create a character for themselves once they register for the game. They will be able to choose the class, race, and also gender for their character. Every novice starts the game with *Level 1*, and their levels increase as they gain more experience points, Experience Points. The player improves in levels when their character reaches a certain Experience Points. As the level goes up, their character will have better attributes, more advanced skills, become more resistant and will have greater choices for more powerful equipment. There are three different points in the game, Experience Points, Battle Points, and Health Points to measure the overall quality of the players.

Each player has various quests to accomplish along with their journey in the game. First of all, there will be the Player versus Environment Combat (PvE) in which players will have battle with monsters in the virtual world in order to become stronger and increase their level. In addition, players will be given freedom of choice between the Player versus Player Duel (PvP) and conversations on different topics as a challenge. If players choose a PVP duel, the winner will be rewarded with Experience Points, Battle Points (for purchasing some unique and powerful items they usually cannot buy with gold coins), and also Gold Coins; whereas the failed one will be debuffed so their movement speed and vitality will decrease dramatically by 70%. To recover their original game stats, players should chat with other players about a topic under a time limit (usually 10-15 minutes), otherwise, they have to wait for two hours to be recovered automatically. However, if players opt for having conversations on different topics voluntarily rather than having a PVP duel, all participants of the conversation will be rewarded with the same amount of Experience Points and Gold Coins as the winner of PVP duel, but obtain relatively less Battle Points. Players will be given clear guidance on how to move freely in the virtual world. For instance, giving them instructions on how to use different keypads for walking, running, and jumping etc. We will give them absolute direction commands, such as North, northeast, west, and so on. Besides, contextual and landmark directions are also provided for them as reference points and help them identify their current location and target place. We will have an in-game virtual economy in the game. Players will have a welcome bonus as a starting point and earn more gold coins as they are involved in playing. Gold coins can be earned in different ways: if the players win in a PVE combat and kill a monster, they will get gold coins as a reward or a loot from the monster; players can also gain equipment as loots from monsters so they can use these loot equipment or sell it for gold coins. Players accumulate gold coins and buy more advanced equipment in order to achieve more success and improve their levels in the game. A frail economy leads to a very few purchase options so that the players will not be able to buy more powerful equipment. As a result, there

will be less chance for them to win in a battle. Apart from gold coins, the Battle Points can serve as another in-game currency. There will be some unique powerful equipment which are purchasable only with Battle Points, so the players will be motivated to have more PVP duel since Battle Points are given more if they win in PVP duel. In result, it will facilitate conversations among players since the failed players from PVP duel are asked to take part in a conversation.

5. Conclusion

We have presented the ultimate goal of our project and proposed a new method for collecting dialogue data for a low resource language, Uyghur. We have introduced the initial design of this method, namely via implementing games with a purpose (GWAPs) for collecting naturally produced dialogues from the game players. We plan to design and implement a MMORPG using the RPG Maker MV Game Engine in which the players will have the ability and the freedom to explore the virtual world according to their levels, and they are given opportunities to accomplish some missions or challenges during the game such as Player versus Environment Combat (PvE), Player versus Player Duel (PvE), and conversations on different topics. Apart from this, players who fail in a PvP battle can recover their original game stats by chatting with others in the chatroom, otherwise, they will be waiting for two hours to be recovered automatically. In this way, we can encourage and motivate players to participate in the conversations. We also plan to integrate the Dialogue Experimental Toolkit (DiET) with the chatroom system of the game so that we can have some control over the interactions of participants, and can improve the data collection process as well. At the same time we can also save our time and costs for transcribing and post-processing the data with DiET.

We will continuously improve the design of the game and start implementing it with the RPG Maker MV Game Engine. After completing the first version of the game, we will test it by recruiting some players and get feedback from them for further improvements.

References

- Bartle, R. A. (2004). *Designing virtual worlds*. New Riders.
- Carletta, J., Garrod, S., and Fraser-Krauss, H. (2000). Placement of authority and communication pattern in workplace groups. *Small Group Research*, 29:531–559.
- Cooper, S., Khatib, F., Treuille, A., Barbero, J., Lee, J., Beenen, M., Leaver-Fay, A., Baker, D., Popović, Z., et al. (2010). Predicting protein structures with a multiplayer online game. *Nature*, 466(7307):756–760.
- Eshghi, A. and Healey, P. G. (2016). Collective contexts in conversation: Grounding by proxy. *Cognitive science*, 40(2):299–324.
- Ginzburg, J. and Fernández, R. (2005). Scaling up to multilogue: some benchmarks and principles. In *Proceedings of the 43rd Meeting of the Association for Computational Linguistics*, pages 231–238, Michigan.

- Ginzburg, J., Yusupujiang, Z., Li, C., Ren, K., and Łupkowski, P. (2019). Characterizing the response space of questions: a corpus study for english and polish. In *Proceedings of the 20th Annual SIGdial Meeting on Discourse and Dialogue*, pages 320–330.
- Ginzburg, J. (2012). *The interactive stance*. Oxford University Press.
- Healey, P., Purver, M., King, J., Ginzburg, J., and Mills, G. (2003). Experimenting with clarification in dialogue. In R. Alterman et al., editors, *Proceedings of the 25th Annual Conference of the Cognitive Science Society*, pages 539–544. Mahwah, N.J.: LEA.
- Khatib, F., DiMaio, F., Cooper, S., Kazmierczyk, M., Gilski, M., Krzywda, S., Zabranska, H., Pichova, I., Thompson, J., Popović, Z., et al. (2011). Crystal structure of a monomeric retroviral protease solved by protein folding game players. *Nature structural & molecular biology*, 18(10):1175–1177.
- Łupkowski, P. and Ginzburg, J. (2016). Query responses. *Journal of Language Modelling Vol*, 4(2):245–292.
- Poesio, M., Chamberlain, J., Kruschwitz, U., Robaldo, L., and Ducceschi, L. (2013). Phrase detectives: Utilizing collective intelligence for internet-scale language resource creation. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, 3(1):1–44.
- Von Ahn, L. and Dabbish, L. (2004). Labeling images with a computer game. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 319–326.
- Von Ahn, L. and Dabbish, L. (2008). Designing games with a purpose. *Communications of the ACM*, 51(8):58–67.
- Von Ahn, L., Liu, R., and Blum, M. (2006). Peekaboom: a game for locating objects in images. In *Proceedings of the SIGCHI conference on Human Factors in computing systems*, pages 55–64.