

# A Cross-lingual Messenger with Keyword Searchable Phrases for the Travel Domain

Shehroze Khan, Jihyun Kim, Tarik Zulfikarpasic, Peter Chen, Nizar Habash

New York University Abu Dhabi

{shehroze, jihyun, tz638, peter.wei.chen, nizar.habash}@nyu.edu

## Abstract

We present Qutr (Query Translator), a smart cross-lingual communication application for the travel domain. Qutr is a real-time messaging app that automatically translates conversations while supporting keyword-to-sentence matching. Qutr relies on querying a database that holds commonly used pre-translated travel-domain phrases and phrase templates in different languages with the use of keywords. The query matching supports paraphrases, incomplete keywords and some input spelling errors. The application addresses common cross-lingual communication issues such as translation accuracy, speed, privacy, and personalization.

## 1 Introduction

The increasing mobility of today's world population naturally leads to the need for easier and more accurate cross-lingual communication tools. Solutions to help transcend language barriers range from using body language, travel phrasebooks and, more recently, machine translation software and hardware devices. However, these methods come with different limitations. Traditional travel phrasebooks, while useful for looking up travel-specific vocabulary, require tedious effort to flip through pages and only allows one-way communication. Recent advancements in machine translation, such as using Google's neural machine translation (Wu et al., 2016), has enabled fast, accurate translations between numerous language pairs, but has limitations for travelers since most machine translation systems are trained on the news domain, and are not optimized for conversational or travel domain language. In this demo paper, we present Qutr (Query Translator), a real-time cross-lingual messaging app. Qutr relies on querying a database that holds commonly used pre-translated travel-domain phrases with the use of keywords. While both users need to have the app downloaded, each user only needs the database for their language of preference, since Qutr internally uses an interlingual representation. As such, a Qutr user need not know what the other user's language is. Qutr addresses the two limitations mentioned above by providing an easily searchable phrasebook that focuses on travel domain conversational language.

We next present some related work (Section 2), followed by system description (Section 3). After concluding and discussing future directions (Section 4), we summarize the demo plan (Section 5).

## 2 Related Work

The mobile application Yochina aimed to aid cross-lingual and cross-cultural understanding by supporting dialogues between English and Chinese, and German and Chinese (Xu et al., 2014). This app addressed many of the problems of existing apps for cross-lingual communication in travel, but still faced issues of operating on a single device and relying on a more manual input of phrase selection, which reduced the speed of communication. Finch et al. (2011) suggest the use of picture icons as input sequences to speed up the communication process. Other translation apps, such as Speak&Translate, support bidirectional and speech-to-text conversation, but all require sharing of a single device, which travelers might find uncomfortable when communicating with a stranger in a completely new environment. Furthermore, apps that rely on speech-to-text interfaces suffer from reduced conversational privacy, lowered quality in noisy environments and diminished accessibility for deaf and hard-of-hearing individuals.

---

This work is licensed under a Creative Commons Attribution 4.0 International License. License details: <http://creativecommons.org/licenses/by/4.0/>

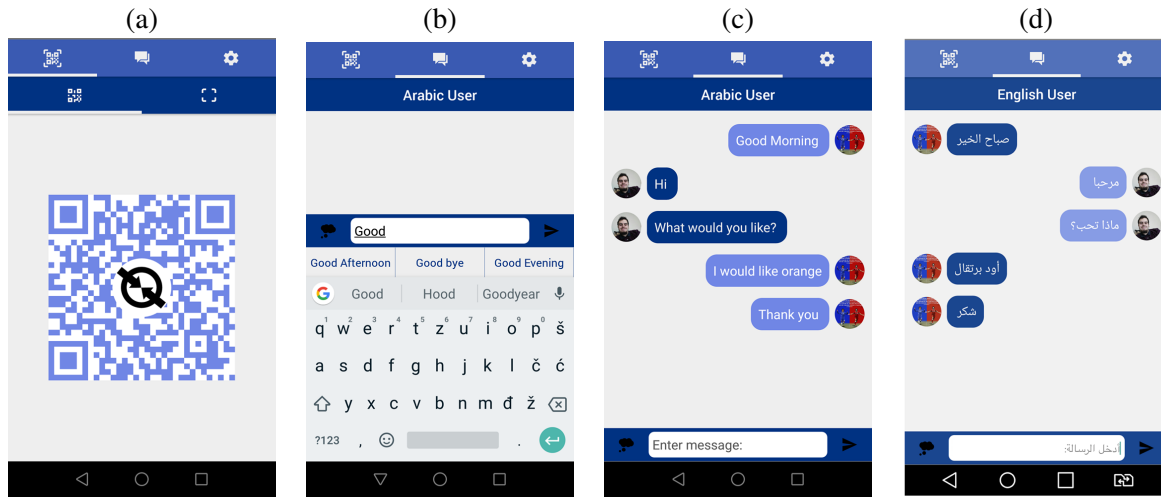


Figure 1: Qutr User Interface screens: (a) QR code, (b) Phrase Suggestions (c) English user side conversation, (d) Arabic user side conversation

### 3 System Description

Our system, Qutr, tackles the abovementioned issues with a real-time messenger app that enables users to communicate on their individual devices in a fast, highly-accurate and user friendly manner. Qutr models a modern form of a travel phrasebook that is easily searchable by natural language querying and that operates on separate devices. Our phrasal database is queried with users’ keywords in order to retrieve suggestions for best matches, which are then matched to existing, high-quality translations. Unlike other translation apps that require databases for the source and target languages to be pre-downloaded, Qutr only expects databases for the individual user’s language, which relate it to the Qutr interlingua representation. This allows one Qutr user to communicate with any other user regardless of their preferred language. While the prerequisite of having the app installed on both users’ end could be a potential drawback, it addresses the discomfort and risks of sharing a single device. Qutr users establish instant connection using QR code scanning, thus preventing delay or disruption in the flow of real-life communication. The QR code ‘handshake’ supports privacy as the two users do not need to share any information about themselves. The text input and output interface also supports privacy and overcomes the challenges faced when one or both users are deaf or hard-of-hearing. To maximize the usability of Qutr, we have built fully functioning versions in both Android and iOS. We imagine that Qutr can be first rolled out to airport information desks and hotel front desks where travelers are likely to use it more. Next, we present the various Qutr system components.

#### 3.1 User Interface

Our application interface follows a very intuitive schema. Upon opening the application, users are directed to a screen displaying their unique QR code which they will scan to setup a connection with the other user, as shown in Figure 1 (a). When one of the users scans the other’s QR code, they are both redirected to a private chat screen. They begin by inputting keywords and selecting one of the shown suggestions in the suggestion bar that best matches what they wish to communicate – Figure 1 (b). Users have to select from at least one of the suggestions (or concepts) before sending it, ensuring that the translated version exists on the other side. A sample conversation between an Arabic-speaking user and an English-speaking user is shown in Figure 1 (c,d).

#### 3.2 Phrase Corpus Design

The phrases that users select and communicate with are queried from our pre-translated phrase database. The current version contains translation triples of 3,500 iconic phrases in Arabic, Chinese and English. Each of these triples is associated with a Concept ID and a Phrase Structure tag. The *Concept ID* is a



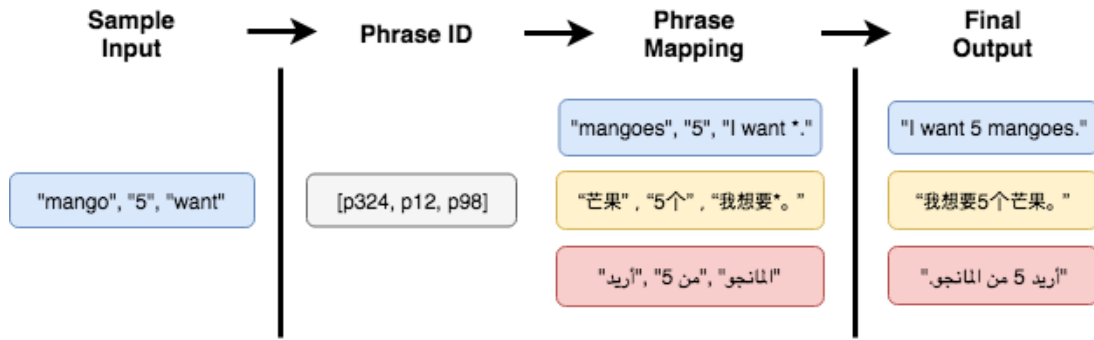


Figure 3: Sentence Generation Process

language questions from user input queries has many useful applications, not only for web search engines, question answering systems and bot communication systems, but also for digital assistants where guiding users to their exact needs is paramount. While they reduce this query-to-question conversion to a translation problem, we utilize the same idea to navigate our users' conversations through suggesting phrases from our database that match an input query.

Ortega et al. (2003) suggest the use of a trie to rapidly search autocompletion strings from user input queries; autocompletion terms stored along the trie are dynamically scored higher if they match user input every time they enter a new character. The authors encourage tagging these terms, enabling efficient searching due to the added filtering and contextualizing. Following a similar implementation scheme, we stored our phrases (iconic and paraphrases) in a trie, where each of the constituent words in a particular phrase are stored with the corresponding concept ID of that phrase. Figure 2 depicts a possible English phrase trie. Note how the paraphrase words *Hello*, *Hi*, and *Hey* will all return the same concept ID [p1].

The phrase suggestion algorithm iterates over the input characters, traversing the trie from the root node. If an exact match is found, we return the corresponding list of concepts. Otherwise, we *autocomplete* by traversing from the last reached node to all possible completion nodes. The concepts are ranked with preference for exact match and minimal edit distance from the input word. The trie also tolerates one character error per word. Imperfect word matches and autocompletions are ranked lower in the identified set of phrases. If users input custom words that are not in our database, we send the words themselves as foreign words. Digit sequences and emoji are also sent as is.

### 3.4 Paraphrase to Iconic Phrase Mapping

It is not always possible to have one-to-one mapping of paraphrases between languages since many colloquial expressions are language and culture specific. We therefore handle paraphrases at language-specific levels by adding extra information to the concept IDs of each paraphrase. For example, "Hello" may have a concept ID of [p1] while "Hey," its paraphrase, would be linked to [p1.en.00] to indicate that it is the first paraphrase of concept [p1] in our database. This additional paraphrase marking scheme can easily scale not only to cover synonyms, but also gender and other inflections of an iconic phrase. For example, in Arabic, the phrase "Would you like coffee?" would have two forms depending on gender of the second person. In this case, we would append extra information to its concept ID to resemble [p500.ar.00.sf] and [p500.ar.00.sm] to differentiate between singular female and singular male versions of a phrase depending on the other user's profile. This is a possible solution for correctly personalizing machine translated conversations based on the user's profile-specified gender, which no existing translation systems do to our knowledge. The current prototype version does not yet support translations that account for inflections in our existing phrases, but we plan on adding it in the future.

### 3.5 Cross-lingual Sentence Generation

Sentence generation in Qutr happens twice per message: once to the user's source language (from selected concepts) and once to the target language (from received concepts). The former serves to map into the iconic phrase from the query which may have a typo or is a non-iconic paraphrase. The generation

function takes in an unordered array of concept IDs and target language as two parameters and returns the full sentence as an output (Figure 3). The algorithm relies on a set of language-agnostic generation rules that order phrases by looking at the phrase structure tags of each phrase. The parallel phrase translations, which maintain the same syntactic roles and template structure for each concept ID, guarantee consistency and universal applicability of the rules. The nature of task-driven communication between different native speakers also means that the *sentences* can be very succinct, informal or even incomplete.

## 4 Conclusion and Future Work

We presented Qutr, a cross-lingual messenger app for the travel domain with support for English, Chinese and Arabic. Qutr’s interface relies on users in one language utilizing query words to find and construct phrases that are sent as concept IDs to users speaking a different language. In the future, we plan to further improve on coverage and conversation flow. We plan to add not only more iconic phrases, but also paraphrases. We plan to complete gender inflections for existing phrases to fully support the correct inflection in conversations. Furthermore, while we have conducted limited pilot tests, we plan a large multi-user evaluation to test the quality of Qutr compared to using phrasebooks or Google Translate. We also plan to add new language phrase databases for wider multilingual support. With the addition of more languages, we envision that Qutr could be used as a language learning tool by giving users an option to view the translated versions of the phrases they wish to communicate on the other side.

## 5 Demo Plan

In the demo of this work, we will provide six mobile phones (three iOS and three Android), with two phones for each of the three supported languages: Arabic, English and Chinese. We will also provide a laptop showing an instructional video demonstrating how to use the Qutr app, and provide some specific scenarios to guide conference attendees interested in this app to role play use cases with the demo presenters and with each other. An example of the demo video is shown here: <https://vimeo.com/262015667>.

## Acknowledgments

This effort has been supported by the New York University Abu Dhabi Science Capstone Project Fund. We would like to thank Ms. Ella Noll and Professor Sebti Foufou for helpful discussions.

## References

- Andrew Finch, Wei Song, Kumiko Tanaka-Ishii, and Eiichiro Sumita. 2011. picoTrans: Using pictures as input for machine translation on mobile devices. In *Proceedings of the International Joint Conference on Artificial Intelligence*.
- Adarsh Kumar, Sandipan Dandapat, and Sushil Chordia. 2018. Translating web search queries into natural language questions. In *Proceedings of the Language Resources and Evaluation Conference, Miyazaki, Japan*.
- Ruben E Ortega, John W Avery, and Robert Frederick. 2003. Search query autocompletion. US Patent 6,564,213.
- Slav Petrov, Dipanjan Das, and Ryan McDonald. 2012. A universal part-of-speech tagset. In *Proceedings of the Language Resources and Evaluation Conference, Istanbul, Turkey*.
- Junyi Sun. 2017. Jieba - Chinese text segmentation module.
- Yonghui Wu, Mike Schuster, Zhifeng Chen, Quoc V Le, Mohammad Norouzi, Wolfgang Macherey, Maxim Krikun, Yuan Cao, Qin Gao, Klaus Macherey, et al. 2016. Google’s neural machine translation system: Bridging the gap between human and machine translation. *arXiv preprint arXiv:1609.08144*.
- Feiyu Xu, Sven Schmeier, Renlong Ai, and Hans Uszkoreit. 2014. Yochina: Mobile multimedia and multimodal crosslingual dialogue system. In *Natural Interaction with Robots, Knowbots and Smartphones*, pages 51–57. Springer.