# How Do People Interact in Conversational Speech-Only Search Tasks: A Preliminary Analysis
### Codebook

J. R. Trippas

Monday 27$^{\text{th}}$ February, 2017

## Contents

# 1 Introduction

This codebook is part of the paper *How Do People Interact in Conversational Speech-Only Search Tasks: A Preliminary Analysis* [2]. This document describes the codes for annotating spoken conversational search interactions. It outlines important characteristics of the identified utterance or turn. The codes presented in the paper are from a preliminary analysis and therefore not definitive. In our further analysis we expand and refine these preliminary codes on the complete data set which will be soon released.

Thus, we expect that the annotation scheme would be refined depending on the area of interest. By accepting these codes we increase the likelihood that annotations are reusable and transferable in other projects and for other purposes.

This annotation schema is relevant to our research aims and was designed after inspecting the selected data; the data set was then coded using these data-derived codes [1]. The codes are generated on the recordings and transcriptions. We are not able to share those recordings, however the transcriptions are generated verbatim (i.e., word for word, exactly as said) and are a good representation of those recordings. Pauzes are transcribed as "...". Notes are added in order to understand the coding where possible.

## 1.1 Visual Representation of Codes

The following representation provides a visual overview of the coding schema. In this schema from left to right is how the conversation develops. All even turns are utterances by Retriever, all uneven turns are utterances by User.

Turn 1: The first turn from that particular search task called Information Request (also referred to as Initial Information Request).

Turn 2: Response from Retriever to User's Turn 1 from any of the following themes: *Meta-communcation, SERP or Scanning Document.*

Turn 3: In this preliminary analysis only turns that respond to *Meta-communication Theme* are presented.
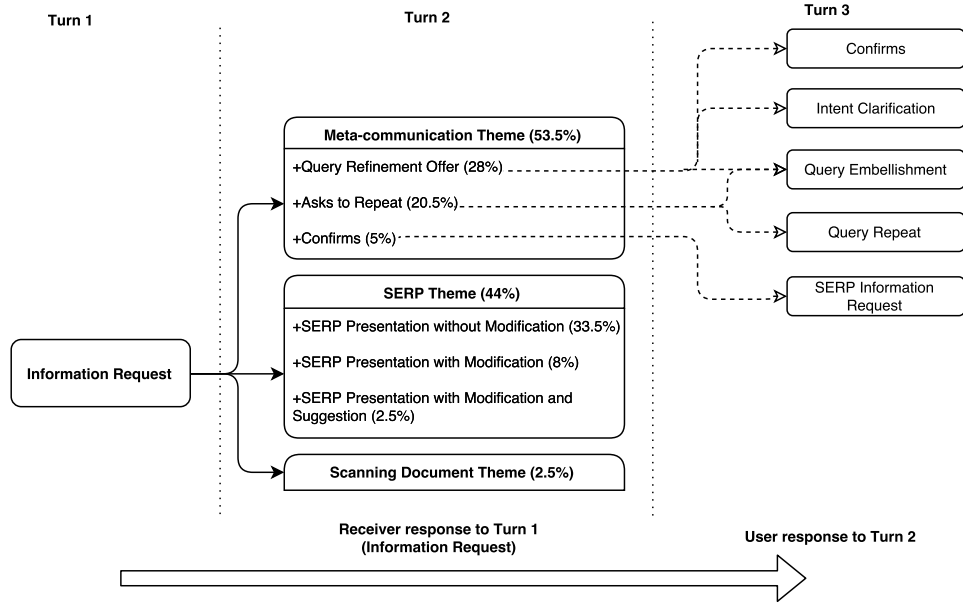
Figure 1: Interaction Theme Map (First three turns).

# 2 Codes for Both User (A) and Retriever (B)

This section covers codes which are used by both User (A) and Retriever (B). These codes are generic codes or utterances.

## 2.1 Confirms

Both User and Retriever is the Confirm code such as:

- `yeah`

- `Yes yes`

- `yea`

# 3 Codes User (A)

This section covers the codes which are generated for User (A). The codes are ordered alphabetically.

### 3.1 A:Initial Information Request

The Initial Information Request is the first information request User pose after reading the backstory. No information has been transferred between User and Retriever about the backstory until User initiates the Initial Information Request. User can start the Initial Information Request with a command-like utterance such as `Can you type in...` or by posing query words.

Initial Information Requests are posed in **query format** such as:

- `In which countries...  in which european countries...  do they grow cinnamon`

- `Can you type in...  effectiveness of new security measures at airports`

- `new security measures at airports`

### 3.2 A:Intent Clarification

User specifies or clarifies their Information Request and is different to **A:Query Embellishment** (see Section 3.3) meaning that the utterance is less query focussed and resembles more of a natural language utterance.

### 3.3 A:Query Embellishment

User enriches the given query for example by adding or removing query words. The query embellishment code is used by User with having a good understanding of what the query was to be able to embellish the query.

An example of this code:

- `effectiveness...  new security measures at airports`

- `US jobs that have been outsourced to india`

- `sure it's health and algae A L G A E and seaweed and kelp`

### 3.4 A:Query Repeat

User repeats the Information Request. The repetition is not altered in any way except that the repetition of the Information request can be in whole or in part.

### 3.5 A:SERP Information Request

User requests more information about the SERP page, this can be related to the position of a search result, the overall gist of the SERP page, or asking if a specific result is displayed.

- `What results came up?`

- `so what's the first result that you get...  do you get a...`

## 4 Codes Retriever (B)

This section covers the codes which are generated for Retriever (B). The codes are ordered alphabetically and not by theme.

### 4.1 B:Asks to Repeat (Meta-communication Theme)

Retriever access Asks to Repeat code with utterances which explicitly ask to repeat the previous utterance such as:

- `Can you repeat the sentence?`

- `Sorry?`

- `Pardon`

The Ask to Repeat code is also accessed by implicitly asking to repeat the utterance with repeating partially what has been said. Implying that the other party repeats the utterance.

### 4.2 B:Query Refinement Offer (Meta-communication Theme)

Retrievers present a possible refinement for an information request. A query refinement offer can be posed by repeating the full information request together with adding some query words or by asking to alter the query in a specific way.

- `international airport brussels bombing...  do we want to keep all of that?`

- `Can I put in United States?`

- `so we're looking for a particular brand?`

Retriever can also communicate with User that they are *Reformulation* the query.

```
so instead of effectiveness improvements
```

or ask specific questions allowing User to respond back:

```
is there a particular country you are after
```

## 4.3   B:Scanning Document Theme

Retriever presents a document to User.

## 4.4   B:SERP Without Modification (SERP Theme)

Retriever reads out the text of the SERP as the text has been written. Retrievers have control over the following aspects of the SERP:

- Only read the title or parts of the title;

- Only read the snippets or parts of the snippet;

- Read out the title or snippet in different ranked position than presented in the SERP;

- Number the search results (eg. "The first result is...");

- Read out URL

Retrievers **do not** interpret what has been written in the SERP in order to present the information to User.

## 4.5   B:SERP Without Modification and Suggestion (SERP Theme)

Retriever presents the written document with alterations as presented in Section 4.4 and makes a suggestion which move could be next on or how to specify the search.

An example of this code is

- ```
  I have scholar articles as well as alcohol consumption per
  capita...  would you like to see more of that
  ```

## 4.6  B:SERP With Modification (SERP Theme)

Retriever does not reads out the text of the Search Engine Result Page as the text has been written, instead Retriever customises what they read from the SERP such as synthesising or giving an overview.

Thus Retrievers form an interpretation of what they read and present this interpretation to User.

- In this example Retriever is giving an overview of the page and what the most prominent features on this SERP are: `uhm...  yeah uhm... it pops up...  it talks about the aqueduct and stuff and... yeah...  pretty much more based on it's history`

- Retriever interprets images in the Google images SERP: `they use it for chair furniture uhm...  and then for uhm...  as a pot.`

- Retriever says the following to User about a search engine result: `well uhm...  the first result is just a company`. However, the result does not mention anything related about a company but Retriever has made an interpretation of the first search result.

**Note:** This code was frequently given by comparing the transcripts to the screen capture software allowing us to detect the modifications from the SERP.

# 5  Outstanding Issues

- Release transcribing protocol
- Release complete data set
- Release complete code book

# References

[1] V. Braun and V. Clarke. *Successful Qualitative Research: A Practical Guide for Beginners*. Sage, 2013.

[2] J. R. Trippas, D. Spina, L. Cavedon, and M. Sanderson. How do people interact in conversational speech-only search tasks: A preliminary analysis. In *Proceedings of the 2017 ACM on Conference on Human Information Interaction and Retrieval (CHIIR)*. ACM, 2017.