The VoiceOver Design for Tables

Ruochen Kong

1 Overview

The internet occupies almost all niches of our daily lives. In order to keep up with trends and not miss updated information, visually impaired people are obliged to use computers, especially touch-enabled tablets. Prior voiceover projects have already successfully improved the accessibility of visually impaired people, but a method to efficiently and effectively represent a table is still unclear. In this report, I aim to design a voiceover technique to represent nested tables. The technique is designed to meet the basic demands of reading tables and provided with an evaluation process afterward. The rest of the paper is organized as follows. Section 2 presents the analysis on the demands of reading tables and general concerns of current voiceover strategies. Section 3 provides the main design. Section 4 introduces the process of evaluation.

2 Analysis

A brief analysis of user requirements is a major step before starting designing. In the context of this report, I separated the analysis into two parts, (i) what information would the users want to receive from the table and (ii) what specific care or functions do they need.

2.1 Properties of tables

Tables are normally used to provide a better understanding of too detailed or complicated information [4]. When reading a table, the readers would expect to quickly see the values and to compare relative contents [4, 5]. To satisfy these expectations as much as possible, the design of voiceover should consider a proper way to locate a specific content and to order the contents when dealing with a comparison task. For example, locating content by starting from the top left corner is not considered a proper positioning method. To determine the order, the user may either compare in rows or in columns, which needs to provide an option to the users. To summarize, two questions need to be solved during design:

1. How to locate a specific content?

2. How to order the contents in comparison tasks?

2.2 Concerns of the voiceover

Previous research has already investigated the satisfication of existed voiceover techinque on iphones and ipads. Based on their research, several probelems shoud be considered [1].

1. *The valid functions of some voiceover technique are not clearly announced.*

2. *The gestures of a serious functions may not be in an appropriate order.*

3. *The major function buttons, such as confirm or cancel, may not be easy to access.*

This design will avoid the mentioned problems as much as possible.

3 Design

With the concern of all the announced problems, the design started by set up several detailed goals: (i) design control actions, (ii) instruct functions, (iii) announce the headers, and (iv) locate the contents.
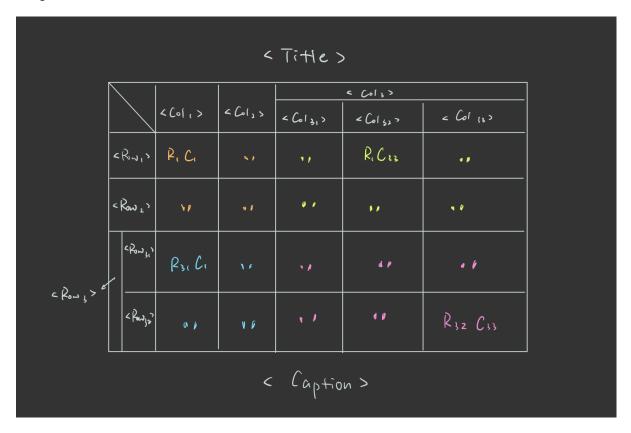

Figure 1: Sample table used during designing

3.1 Controllers

The common options of control gestures including swiping left or right, swiping up or down, single click, double click, and pressing. Hence, oral inputs forms another type of controllers.

1. **Swiping**

   When browsing the table, the swiping direction indicates the change of location in the same direction.

   When the location is next to nested columns or rows, swiping into the nested part will change the location to the most general ones. For example, when locating at R_2, swiping down will goes to R_3; and when locating at C_2, swiping right will goes to C_3.

   When the location is inside nested columns or rows, swiping to detailed contents will follow the top left priority. For example, when locating at R_3, swiping right will goes to R_{31}; and when locating at C_3, swiping down will goes to C_{31}.

   Swiping up twice within 1 sec will move the location to the top of the entire column, and similarly, swiping down twice will move to the bottom, swiping right twice the most right, and swiping left twice will move to the most left.

   When swiping in a middle of a sentence spoken by the voiceover, swiping left or up means to repeat the previous word, and swiping right or down means to skip the current word.

2. **Single Click** State the current location in the table.

3. **Double Click**

   When double clicking in a middle of a sentence spoken by the voiceover, it means to pause and to start again by another double clicking with in 5 seconds.

   Double clicking also signals the finishing of an oral input or the yes answer of a yes or no question.

4. **Pressing**

   During an oral input, pressing for 2 seconds will delete the most current input word, and pressing for 5 seconds will cancel the input. When deleting one word, the voiceover will respond "Delete [the word]"; when deleting all, it will respond "Cancel input".

   Otherwise, pressing for 2 seconds will move the location to the top-left corner.

5. **Oral**

   To start an oral input, users should say "voiceover" and wait for a "Yes" answer, just as Siri. Then several inputs are valid which will be stated in following sections. If an invalid input is received, then the voiceover will respond as "Sorry I don't understand".

3.2   Instructions

In order to clearly announce the way to use this VoiceOver method, instructions should be provided to the users, but they would be annoying when the users are already familiar with the system. Thus, I decided to always provide the instructions at the start, allow users to turn this function off, and create a help list to restart the function or for instruction for a specific function.

Swipe to move the location in the same direction. Double click to confirm or pause. Press 2 seconds to delete one or move to the top-left and press 5 seconds to delete all. Say "Voiceover" to start oral inputs. Valid oral inputs include title, header, indices, content, and locate. The use of "content" and "locate" should follow the rules: (i) say "Row" and/or "Column" before listing the exact names, (ii) double click when finish row names and/or column names, and (iii) say "start" and double click to finish the input. Say "No instruction" to turn off the automatic instruction section, say "Default" to restart, and "help" to repeat the instruction part once.

3.3   Title, Headers and Indices

The titles and caption are always important messages of a table, so they should be the start of a table if available. The actions of voiceover for the example table are:

1. *The title is: [Title]. The caption is: [Caption].* **Oral input:** title.

2. a. *Headers of columns are: [Col_1], [Col_2], [Col_3] with subheaders [Col_{31}], ... , (new header [Col_4]).* **Oral input:** headers / headers for columns.

   b. *Indices of rows are: [Row_1], [Row_2], [Row_3] with subheaders [Row_{31}], [Row_{32}] , (new index [Row_4]).* **Oral input:** indices / indices for rows / headers for rows.

   Oral input: headers and indices.

Inside sentences, ":" corresponds to an 1 second break, "." corresponds to an 1.5 seconds break, and "," corresponds to a 0.5 second break.

3.4 Contents

When the location is at the boundary and the user is swiping out of the boundary, the voice over will say: *"Top / bottom / right / left boundary arrived."*

When single click or move to another location, the voiceover will state the content in the form of:

Content: [content]. In Row [row index], (under super row [super row index]). In Col [column header], (under super column [super column header]).

For "Locate", only singel row and column are allowed. For example, "Column [Col_3]" is invalid, and if it is recieved, the voiceover will respond "nested headers, select a subheader from [Col_{31}], [Col_{32}] and [Col_{33}]."

For "content", multiple inputs of rows and columns are accepted. In the condition with multiple rows and/or columns, the order of output should be decided by the use with an additional question. Here is an example:

User: Content. Row: [Row_2], [Row_{31}]. column: [Col_{31}], [Col_{32}]. Start.

Voiceover: Double click for contents in the same row first. No move for 2 seconds otherwise.

User: [Double click]

Voiceover: In row [Row_2]: [R_2C_{31}] in column [Col_{31}], [R_2C_{32}] in column [Col_{32}].

Voiceover: In row [Row_{31}]: [$R_{31}C_{31}$] in column [Col_{31}], [$R_{31}C_{32}$] in column [Col_{32}].

4 Evaluation

I decided to use heuristic evaluation because tasks assigned by ordinary people would miss the specific demands of visually impaired people. This type of evaluation is not sufficient with a single tester [2], so I decided to recruit 10 testers, which is efficient and would find over 85% of usability problems [2]. Instead of assining the tasks directly, I will provide them with several nested tables and allow them to brows through. I will then ask several questions about (i) degree of user control and freedom, (ii) effectivity of error prevention, (iii) amount of information to remember, (iv) capability to achieve the user goals, and (v) usefulness of help function [3]. I will ask testers to give each question a score between 1 to 5 with a brief rationate. Then generate several plots and understand the plot with the help of their reasons.

References

[1] B. Leporini, M. C. Buzzi, and M. Buzzi. Interacting with mobile devices via voiceover: usability and accessibility issues. In Proceedings of the 24th Australian Computer-Human Interaction Conference, pages 339–348, 2012.

[2] J. Nielsen. How to conduct a heuristic evaluation. Nielsen Norman Group, 1(1):8, 1995.

[3] J. Nielsen. Ten usability heuristics, 2005. URL https://www.nngroup.com/articles/ten-usability-heuristics/.

[4] D. J. Slutsky. The use of tables. Journal of wrist surgery, 3(4):219, 2019. ISSN 2163-3916. doi: https://doi.org/10.1055/s-0034-1395165.

[5] University of Colorado Boulder. When to Use Tables and How to Make Them Accessible to Screen Reader, 2022. URL https://www.colorado.edu/digital-accessibility/2016/09/17/when-use-tables-and-how-make-them-accessible-screen-reader-users.