# Heuristic Evaluation of SpeakEasy



**Problem Statement**

SpeakEasy attempts to solve the issues that come with learning a new language, primarily the lack of support when using the new language in real time situations, by providing on-the-go support with contextual phrase suggestions, dynamic speaking support, and progress tracking.

**List of Heuristic Evaluations**

1. **[H2-4. Consistency and standards ][Severity 2][A, D]**
There is an option for hearing all words except the ones in the transcript. This is not keeping with the standards of the app, and the problem SpeakEasy is trying to solve. Perhaps add sound to transcripts as well.

2. **[H2-2. Match between system and the real world] [Severity 2][A]**
When pronouncing a word (in Google glass), it is assumed in the flow that the user always says the word incorrectly at first. However, logically the user could be correct or incorrect from the get-go. A more logical flow would be to give a choice between user saying it wrong/right from the get-go then having those choices lead to a correction/moving on respectively.

3. **[H2-3. User control and freedom] [Severity 4][A, D]**
In Google glass, once the user starts using the application, he can't get back to instruction video page (or to home for that matter) unless he starts over with the original link to the demo. A home button and appropriate links to the instruction page should be provided so that the user can navigate as he wishes. In addition, it is possible to get into an infinite loops in the Google Glass app. Consider this case: the user pronounces the word correctly, but Google Glass makes a mistake and finds the word pronounced incorrectly. There is actually no way for the user to progress with the conversation with SpeakEasy unless the app believes the user speaks the word correctly.

4.    **[H2-4. Consistency and standards][Severity 3][A]**
The application indicates that 'record and translate' is the verbal stimulus for the app to start recording and translating human speech. However, present on the screen at this time is a microphone button. Usually (on iphones at least) the microphone button is meant to be pushed to start recording. Therefore, the user might instinctively try pushing the microphone button instead of reading the instructions. The microphone button should be removed to eliminate confusion, or the stimulus should be changed to the microphone button.

5.    **[H2-6. Recognition rather than recall] [Severity  2][A, C, D]**
When a user pronounces a word correctly/incorrectly, there is a tick / wrong sign on the screen, respectively. While this sign tells the user if he is correct or incorrect, it does not tell why, and so the user must try to remember how he pronounced the word correctly/incorrectly to know how to say it next time. Instead, to promote recognition rather than recall, the tick/wrong sign could be accompanied by a correct/incorrect phonetic representation of the word so the user can recognize which parts of the word he pronounces correct/incorrectly.

6.    **[H2-8. Aesthetic and minimalist design] [Severity 2][A, C]**
There are many repetitive instructions always present on the page, such as the following instruction: the app will…and exit automatically when it is done". These instructions clutter the screen and distract the user from pertinent info. These instructions should be removed and only be present the first time the app is used by someone / if a user requests it.

7.    **[H2-8. Aesthetic and minimalist design] [Severity 2][A]**
Words on the screen are rather large. An example of this is 'speaking…' written in large letters across the screen when the app is speaking. This screen should be especially minimalist because someone using Google glass would probably not want so much of their vision cluttered with words while trying to interact with a person. Since one of the tasks is to enhance real life conversation, the app should not distract the user too much from his environment. These words on the screen should be made much smaller.

8.    **[H2-7. Flexibility and efficiency of use][Severity 3][A]**
To see other options of phrases one can respond with, the user must scroll down/up/left/right. It is a slow process that only provides one phrase at a time. After using it for a while, it started to seem clunky and I wanted to be able to see more phrases at once to go more quickly. To allow flexibility and efficiency for a expert user, perhaps shortcuts could be added to get more phrases on the screen at once.

**9.     [H2-4. Consistency and standards] [Severity 1][A, D]**

Nothing on the progress screen seems touchable. Since the progress screen is to encourage daily practice and to gamify, the user should be able to manipulate the charts and graphs on the screen. The user should have the ability to choose what kind of practice challenges he/she undertakes.  Usually progress screens have graphs/charts that can be tapped to get more detailed information. To stick with that standard, the same strategy can be applied here.

**10.      [H2-10. Help and documentation] [Severity 2][A, D]**

'Practice' seems like a text message conversation. If it is a text conversation, who is the user talking to? Why are the conversations in different colors - what do the colors mean? Are they talking to the app? These are questions I had when using it. To sort out these questions, some instructions could be provided on how to use 'practice' and in fact, the entire iPhone version could use some instructions similar to how the Google glass version has instructions at the beginning.

**11.     [H2-4. Consistency and standards] [Severity 1 ][A, C, D]**

Regarding the text-message-like 'practice' screen, the colours are not standard android / iOS colours. In fact, the colours are not only non-standard, they also vary. Do the varying colours of the texts mean something? If so, this could be explained in the instructions. If not, they should be kept standard.

**12.     [H2-4. Consistency and standards] [Severity 2 ][A, D]**

In the practice page, there are little speaker signs to hear the words and microphones to speak the words into the text (I assume). However, the words that have already been written still have microphones next to them, indicating the microphones can be clicked to record the text. I assume they cannot be rerecorded since the text thread has moved past that. Therefore to avoid confusion, the microphone button should disappear after the text has been sent.

**13.     [H2-8. Aesthetic and minimalist design] [Severity 1 ][A, C]**

When 'speaking' is written on the screen in the Google glass version, there is also a speaker sign beside it. Since they represent the same thing, we do not need to have them both. 'Speaking' could be removed for a more minimalist design.

**14.     [H2-6 Recognition Rather Than Recall][Severity 1 ][C, D]**

In the mobile prototype, after clicking on a word under "Missed Words" to see the translation, the screen that appears has the translated word at the top, followed by the

remaining mispronounced words. Clicking back takes the user all the way to the home screen. If they forgot what the word was that they clicked, they have to navigate to the Missed Words screen once again, to see what word they had clicked on. Allow the user to see both the missed word and translation on one screen, or let the back button navigate to the Missed Words list rather than home.

**15.  [H2-4 Consistency and Standards] [Severity 0 ][B, D ]**
The titles on the mobile app are different sizes. For example, the titles "SpeakEasy" and "Progress" are noticeably smaller than "Practice," "Transcripts," and "Missed Words." This is very nitpicky, but it was noticeable as soon as I went to one of those title screens.
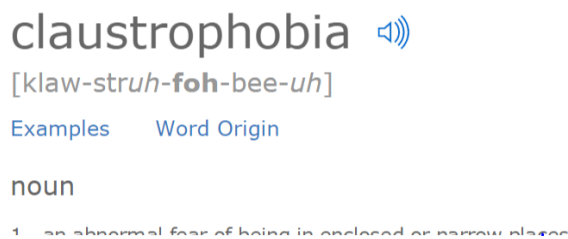Fix: make all the titles the same size (the slightly bigger font looks better in my opinion)

**16.     [H2-2 Match between system and the real world][Severity 3][B, D]**
On the missed words screen, the user should click on the speaker icon to hear the word. However, a two-dimensional speaker button does not indicate to the user to "click me" considering buttons usually have a three-dimensional design. Thus, the system did not "speak my language" by telling me that the speaker was in fact a button to click on. Furthermore, it ignored typical conventions in word pronunciation - Dictionary.com has a good example of a format that assists with learning words (below).
Fix: make it look clickable (three-dimensional)



claustrophobia 🔊
[klaw-str*uh*-**foh**-bee-*uh*]
Examples      Word Origin

noun

1. an abnormal fear of being in enclosed or narrow places

**17.     [H2-2 Match between system and the real world][Severity 2][B]**
In the Transcripts menu, after clicking on "At Restaurant," there is a transcript of a conversation going on in the format of a text. It looks similar to the format of what an iPhone text conversation looks like, but as a lifetime Android user, I was very confused at who was talking when I first came across this screen since the Android layout is significantly different than this layout.
Fix: don't get rid of the layout, but make it extremely clear who said what (maybe add a picture or name to the user's text bubble)

**18.     [H2-7 Flexibility and efficiency of use][Severity 2][B]**
At the start of the Google Glass UI, there is an option to start a tutorial. This would be very annoying for expert users to come across every time and say no to (although I'm not sure if

it actually only appears the first time, since it was not made very clear in the README or handout).
Fix: allow expert users to turn off the tutorial or place it somewhere else for novice's to access

## 19.    [H2-4 Consistency and standards][Severity 1][B]
In the Google Glass UI, instructions for the user and translations for the user to read are both white, potentially making for confusing interactions with the system, especially since this system is so grounded on the fact that it gives users instant feedback and easily readable information to speak out immediately.
Fix: make the instructions a different color or make it look different

## 20.    [H2-4 Consistency and standards][Severity 0][B]
In the Google Glass UI, in the tutorial, the English translation was on the top and the Spanish translation was below that. However, when performing real tasks, this is reversed, with the English translation on the bottom and the Spanish translation on the top.
Fix: pick one translation to be on the top and one to be on the bottom and stick with it (Spanish translations on the top seemed to make more sense for me personally)

## 21.    [H2-8 Aesthetic and minimalist design][Severity 1][B, C]
This violation has to do with the disproportionately large size of the "Missed Words" tab. This single tab has a lot of irrelevant and rarely needed "information" in the sense that there was no need to give users that much room to click on one tab considering it competes with other very relevant units of information(other tabs) and diminishes their relative visibility.
Fix: don't give so much room to click on just a single tab, make all tabs the same size so that the other three tabs don't compete with one single tab for visibility

## 22.    [H2-4 Consistency and standards][Severity 1][B]
The graphs on the "Progress" tab both look the incredibly similar even though one has a positive connotation (New Words) and one has a more negative one (Missed Word). Considering that blue is a color that has been used as a positive color throughout the system (like the checkmark being blue in the Google Glass UI to mean a user did something did something correctly), it struck me as odd that it was the same color used to color the "Missed Words" graph, and it was very glaring when it was right next to the "New Words" graph.
Fix: make the "Missed Words" graph different than the "New Words" graph, consider using different colors for the former as well

## 23.    [H2-6 Recognition rather than recall][Severity 3][B, C]

The user must remember to tap on the screen whenever the user wants to submit something so the system speaks. This action is touched upon in the tutorial, and it's mentioned once when the screen tells the user to "Tap for Phrases," but it requires the user to remember to tap instead of have something indicate when it is appropriate to tap or when a user can tap at all.
Fix: have a small icon of a finger tapping in the corner or some other easy to see yet not invasive icon to let the user know when it is appropriate to tap

## 24.    [H2-8 Aesthetic and minimalist design][Severity 2][B]
One of the screens on the Google Glass UI says "Tap for Phrases." I'm a little confused as to the function of this screen since it immediately follows a screen (the one that translates what the waiter said) that already requires a tap. So, essentially, the user must tap to get to this screen, a screen which has only one function: telling the user to tap to get to the next screen. Perhaps I'm misunderstanding this since it wasn't made too clear in the README or report, but regardless, it seems like a waste to have a screen dedicated to telling the user to tap, even if it is supposed to occur immediately once someone finished talking.
Fix: get rid of this screen, it only adds more work for the user

## 25.    [H2-3 User Control and Freedom][Severity 1][B]
When navigating through the Transcripts tab, if I want to get back to the other tabs while viewing the "October 15 At restaurant" page, I must press back twice instead of there being an easier way to get back to the tabs immediately.
Fix: put a button to take you back to the tabs section

## 26.    [H2-2 Match between system and the real world][Severity 3][B]
In the practice tab, the speaker icons do not look like they should be clicked (same situation as Violation #2). If these icons were not meant to be clicked (since it wasn't made too clear in the report), then this would be an H2-4 Consistency and standards violation since the same icon should be used to do the same function.
Fix: make the icon look clickable (same as Violation #2)

## 27.    [H2-3 User Control and Freedom][Severity 3][B]
In the Google Glass UI, when the app pronounces a word, there is no clear indication about how to cancel this process, which seems really important considering a user may accidentally tap to hear the pronunciation. Since this is an app that prides itself on instant and real-time feedback, it could potentially make situations awkward if long phrases are pronounced and the user has no knowledge of cancelling this during a conversation.
Fix: add a cancellation feature, or if tapping is the cancellation feature, make it clear to the user that tapping will cancel said action

## 28.     [H2-10 Help and documentation][Severity 2][B]

The tutorial does provide all information that the user needs to accomplish a task, but it is not very task focused. For example, the tutorial teaches the user how to record speech last, yet in the prototype's main and functional task, this step is required first, which could potentially be confusing for novices. In this sense, it tells the user information but is not focused on the user's task (which in this case constitutes a specific yet very common scenario).

Fix: run the tutorial like the scenario shown so users are capable of performing the main scenario described in the prototype since it seems like one of the most common tasks the user would use this app for

## 29.     [H2-8 Aesthetic and Minimalist Design][Severity 1][C]

In the Google Glass prototype, the location of the conversation is shown in the screen for the duration of the conversation. After identifying the location, it is not necessary to show it for the remainder of the conversation.

## 30.     [H2-4 Consistency and Standards][Severity 3][C]

In the mobile prototype, under "Practice," the words being spoken have different colors (gray, green, and red), though it is unclear what each of these mean. Add a description of what each of these colors signify.

## 31.     [H2-4 Consistency and Standards[Severity 3][D]

There exists a design mismatch between the Google Glass application and the mobile application. If you removed the title and did not tell me they were part of the same application, I would not be able to tell they were part of the same system.

## 32.     [H2-1 Visibility of System Status][Severity 3][D]

Since the technology will be using Google Glass, there should be some acknowledgment that the two are live and linked. There is no such acknowledgment in the mobile app.

**Summary of Violations**

| Category | # Viol. (sev 0) | # Viol. (sev 1) | # Viol. (sev 2) | # Viol. (sev 3) | # Viol. (sev 4) | # Viol. (total) |
|---|---|---|---|---|---|---|
| [H2-1: Visibility of Status] | 0 | 0 | 0 | 1 | 0 | 1 |
| [H2-2: Match Sys & World] | 0 | 0 | 2 | 2 | 0 | 4 |
| [H2-3: User Control] | 0 | 1 | 0 | 1 | 1 | 3 |
| [H2-4: Consistency] | 2 | 4 | 2 | 3 | 0 | 11 |
| [H2-5: Error Prevention] | 0 | 0 | 0 | 0 | 0 | 0 |
| [H2-6: Recognition not Recall] | 0 | 1 | 1 | 1 | 0 | 3 |
| [H2-7: Efficiency of Use] | 0 | 0 | 1 | 1 | 0 | 2 |
| [H2-8: Minimalist Design] | 0 | 3 | 3 | 0 | 0 | 6 |
| [H2-9: Help Users with Errors] | 0 | 0 | 0 | 0 | 0 | 0 |
| [H2-10: Documentation] | 0 | 0 | 2 | 0 | 0 | 2 |
| *Total Violations by Severity* | 2 | 9 | 11 | 9 | 1 | 32 |

**Evaluation Statistics**

| severity\evaluator | evaluator A | evaluator B | evaluator C | evaluator D |
|---|---|---|---|---|
| level 0 | 0% | 100% | 0% | 0% |
| level 1 | 20% | 26.7% | 33.3% | 20% |
| level 2 | 41.1% | 23.5% | 11.8% | 23.5% |
| level 3 | 18.2% | 36.4% | 18.2% | 27.3% |
| level 4 | 50% | 0% | 0% | 50% |
| **total  (levels 3 & 4)** | 30% | 40% | 20% | 40% |
| **total  (all levels)** | 40.6% | 43.8% | 28.1% | 37.5% |

| evaluator # | # problems found | unique? |
|---|---|---|
| A | 13 | 0 shared with #B -> 13 unique |
| B | 14 | 14 (most found) |
| C | 9 | 7 shared with #A, #B, & #D -> 2 unique |
| D | 12 | 9 shared with #A, #B -> 3 unique |

Graph of this data can be found on next page.