Carnegie Mellon

NATIVE**ACCENT**
Final Report

Matt Guibert | Jon Maloto | Vedant Mehta | Steve Selzer | Josh Weihnacht

Table of Contents

Final Report

Introduction

We are a team of Human-Computer Interaction students at Carnegie Mellon University who have been tasked with evaluating the usability of the NativeAccent software. The purpose of this report is to document our findings on the usability of NativeAccent for the Carnegie Speech development team.

Methods

Our semester-long project used a variety of proven (industry-standard) Usability Evaluation Methods including Contextual Inquiry,[1] Keystroke-Level Modeling,[2] Heuristic Evaluation,[3] Cognitive Walkthrough,[3] and Think-Aloud Usability Studies.[4]

Contextual Inquiry

The first technique we used was Contextual Inquiry, in which we set out to understand the motivations of users through a series of interviews. To set a direction for our interviews, we developed two focus points as a result of a collaboration with the Carnegie Speech development team and Lynda Stucky from ClearlySpeaking. The results of the collaboration are as follows:

Focus 1: Understand how users perceive progress and how feedback influences their performance.
Focus 2: Understand how users interact with the software tool and how efficiently and effectively they accomplish set goals.

We conducted three interviews with new users over the course of 3 weeks. During the interviews, we observed the users performing tasks within the system and asked questions in order to understand the motivations behind their actions. All three interviewees were Indian male university students in their early 20s; while all three were born and raised in India, they were all proficient in English.

Keystroke-Level Model

Next we created a Keystroke-Level Model of the flash activities in NativeAccent. This allowed us to better understand the the complexity of the physical actions required by a user of the site.

---

[1] For information on Contextual Inquiry, please see: Hugh Beyer & Karen Holtzblatt: (1997) *Contextual Design : A Customer-Centered Approach to Systems Designs*. Morgan Kaufman Publishers.

[2] For information on Keystroke-Level Modeling, please see: John, B. E. (2003) Information processing and skilled behavior. . In J. M. Carroll, (Ed.), *Toward a multidisciplinary science of human interaction*. Morgan Kaufman

[3] For information on Heuristic Evaluations and Cognitive Walkthrough please see: Jakob Nielsen & Robert L. Mack: (1994) *Usability Inspection Methods.* John Wiley.

4 For more information on Think-Aloud Usability Studies please see: John, B. E. (1999) Carnegie Technology Education Course #SSD4 User-Centered Design and Testing, Unit 3 Think-aloud Usability Testing

Heuristic Evaluation

We performed a Heuristic Evaluation on the NativeAccent site in which each team member went through NativeAccent looking for issues that conflicted with a standard list of usability heuristics.[5] We then compiled these issues and wrote up the most important findings and included them in the appendixes of the report.

Cognitive Walkthrough

In order to better understand the issues that may be encountered by a new user, a Cognitive Walkthrough was performed on the site. This involved our team going step by step through the tasks of logging in, choosing a new activity, answering a question, and looking at the grades page. At each step of a task we looked for places where a new user would likely encounter an issue.

Think-Aloud Usability Study

Finally, we performed a Think-Aloud Usability Study where we asked a new user to complete a pitch activity, a duration activity, and to view their grades. During each task the user was asked to "think-aloud" so that we could get a better understanding of their thought process.

Usability Issues

There were several central issues that surfaced over the course of our evaluation:

1. lack of an initial overview of the system;
2. lack of understanding of Intelligent Tutor;
3. general navigation issues;
4. feedback on performance;
5. difficulty in understanding system status;
6. visual complexity of flash activities interface.

Lack of an Initial Overview of the System

While users could generally understand *how to use the system*, we observed issues with users understanding *what they should be doing on the system*. For example, users tended to not understand the importance of completing the assessment before exploring practice lessons. Users were also confused by the fact that they were graded when they thought they were just practicing. In an extreme case, one user spent an entire hour looking for a tutorial that could provide him with an overview of the system. This issue may be less prevalent for students who are using NativeAccent as part of a class and have been given an overview by their teacher. However, it was a common issue for students using the site on their own.

Lack of Understanding of Intelligent Tutor

The Intelligent Tutor, which ideally should provide guidance through the system was poorly understood by most users. None of the users understood that it provided a suggested lesson plan or that the suggested order was customized for them. Part of the confusion may have been due to the fact that lessons are ordered differently in the Intelligent Tutor compared with the class home page and the jump menu. One user even thought the Intelligent Tutor was supposed to be a tutorial.

General Navigation Issues

Navigation was an issue for all users and came up in multiple Usability Evaluation Methods. The jump menu was poorly understood and is a non-standard interface element; in particular the jump menu's back and forward buttons were confused with the browser's back and forward buttons. When a user completed a

[5] http://www.useit.com/papers/heuristic/heuristic_list.html

Flash activity, it was confusing for them to have to use the jump menu to get to the next activity. Labels on the breadcrumbs, such as “NA-CMU” and “CS-Web” are unlikely to be understood and the breadcrumbs were often ignored by users.

Feedback on Performance

It was often difficult for users to interpret the feedback from the system. Users were confused about how to read the pitch and duration graphs and were sometimes unsure what terms like “stress” meant. The mid-sagittal diagrams were hard to read. The grades page is also difficult to understand.

Difficulty in Understanding System Status

We observed several issues with the ability to understand the current state of the system. When a question is answered correctly, feedback that the user was correct is minimal and the system quickly skips to the next question. Thus users sometimes got confused and thought that they were still on the previous question after an auto-advance. System error messages were also confusing at times, such as the “Moodle service error”. Other error messages left the user unsure what to do next, for example it is unclear how the “Speech does not match” error differs from simply making mistake in pronunciation.

Visual Complexity of Flash Activities Interface

Finally, there was a class of issue related to the complexity of the interface for flash activities. Users often did not notice the model speaker button until they had used the system for awhile. Even after they noticed it, some users seemed to forget about it later. One user would also forget to hit the record button before answering a question. Users also tended not to notice the clock indicating how much time they had left until after they ran out of time. Finally, the submit button can be difficult for new users to find.

Recommendations

Our team has developed the following recommendations, each of which responds to the usability issues described in the previous section:

1. provide an initial overview of the system;
2. re-brand Intelligent Tutor as Lesson Planner;
3. provide robust navigation that follows web conventions;
4. clarify feedback on performance;
5. provide clear indications of system status;
6. simplify flash activities interfaces;

We will also provide additional recommendations that related to other general issues.

Provide an Initial Overview of the System

The data collected from our Contextual Inquiries suggests that an initial overview that orients users would help with an understanding of the overall service provided by NativeAccent. While there is a tutorial that teaches users how to use the controls in the flash activity, there is no corresponding overview of the user’s workflow within the system. Students without the benefits of a classroom and teacher would require this context in order to engage the system in a productive, edifying manner.

Re-brand Intelligent Tutor as Lesson Planner

The name “Lesson Planner” clearly conveys what the functionality sets out to do: plan the order in which lessons should be completed. In addition, it would be helpful to provide an explanation of how the Lesson Planner and “create new path” functionality work and why it is helpful to follow the lesson plan.

Because of the importance of following the planner, NativeAccent should integrate the functionality into the entire experience by presenting it throughout the system.

Provide Robust Navigation that Follows Web Conventions

Current web conventions suggest that a menu-driven navigation system would improve usability significantly. We suggest that you replace the Jump-To functionality with a navigation menu at the top of the page. The menu options should intuitively group parts of the site, and while we don’t have strong evidence for a specific grouping, below are a few suggestions:

- have a “Home” option;
- have a “Suggested Lesson” option for reinforcing the importance of the Lesson Planner;
- have separate options for the flash activities categories Word Stress (pitch and duration), Consonants (e.g. bank, dog, food, etc.), and Vowels (e.g. out, eye, end, etc.);
- have a “Grades” option for quick access to reports and feedback on performance.

It may be helpful to keep the breadcrumbs, which will help users identify where they are in the site, but in order to improve usability, the labels need to be worded more intuitively.

Clarify Feedback on Performance

The feedback provided when a user answers a question incorrectly needs to be articulated more clearly and graphically represented more intuitively. The data suggests that users would like to see more dramatic feedback when answering a question correctly, like a notification as simple as “Great!”. Because the system quickly auto-advances, it is important that this feedback “pop-out” at the user quickly so that it is perceived before the next question appears.

In addition, the suggestion text should be more interactive, enabling the user to learn from the suggestions by linking each phoneme to a sound byte that, when clicked, plays that particular pronunciation.

Furthermore, the “Grades" page should be redesigned to convey the information more intuitively. While our data does not support specific design changes, the ideas below are motivated by the frequency of general confusion exhibited by users while trying to interpret grades:

- replace the pair of graphs for each activity with a single graph that indicates the current score; show the original score as a line on this graph;
- try consolidating the scores into a single composite score, perhaps a percentage;
- try orienting the data as a horizontal bar graph to more efficiently use space;
- indicate how the user is improving over time by incorporating line graphs of progress over time;
- provide an explanation of how to interpret the graph;
- provide old sound clips to hear original recordings from the first assessment and compare to current clips.

Provide Clear Indications of System Status

While the system currently provides a “Status” bar in the flash activities interface, the placement and visual cues are not obvious to the user, as suggested by the data collected in our Contextual Inquiries and Think-Aloud Usability Study. Many users do not seem to even notice that this status bar exists. A more obvious

visual cue that the system is processing information could be an animated spinning icon in the content field where the user is likely to be focusing his/her attention.

Simplify Flash Activities Interface

Many of the following suggestions are demonstrated visually in our screenshots in Appendix A:

- make model speaker button more prominent;
- make record button more prominent;
- make the positive feedback more prominent;
- dedicate an area for feedback, rather than providing feedback via pop-up;
- prompt with a link to “view grades” when user completes an activity;
- provide a “next lesson” option when user completes an activity;
- make the progress bar more prominent; differentiate between correct, incorrect, and incomplete questions;
- the progress bar could allow users to click on the box for an for a question for faster navigation.

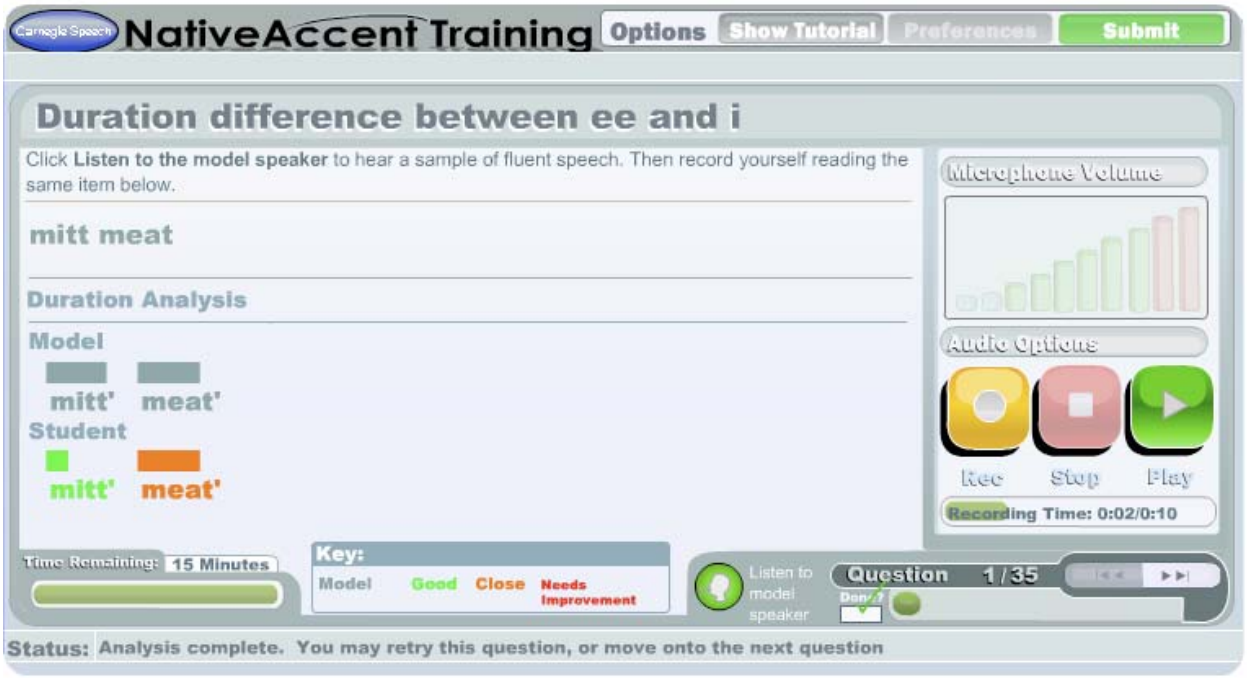


Figure 1. Current NativeAccent interface analyzed by this project.

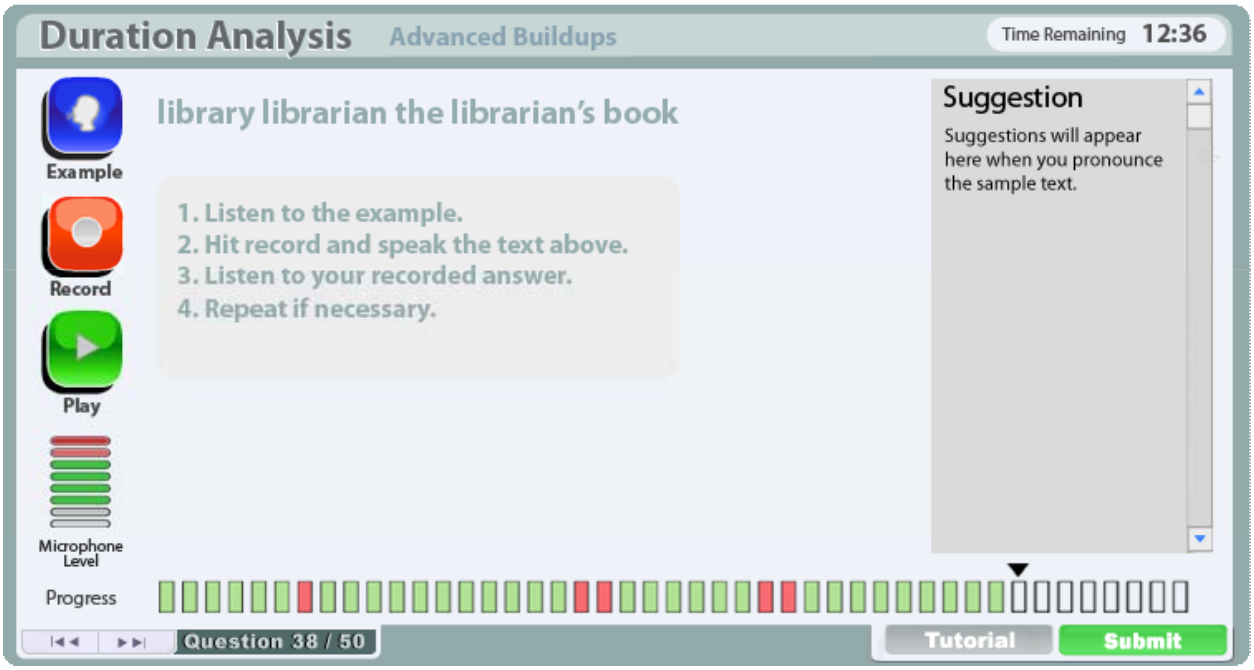


Figure 2. Proposed redesign. Start of activity. Instructions appear before the first question is answered.

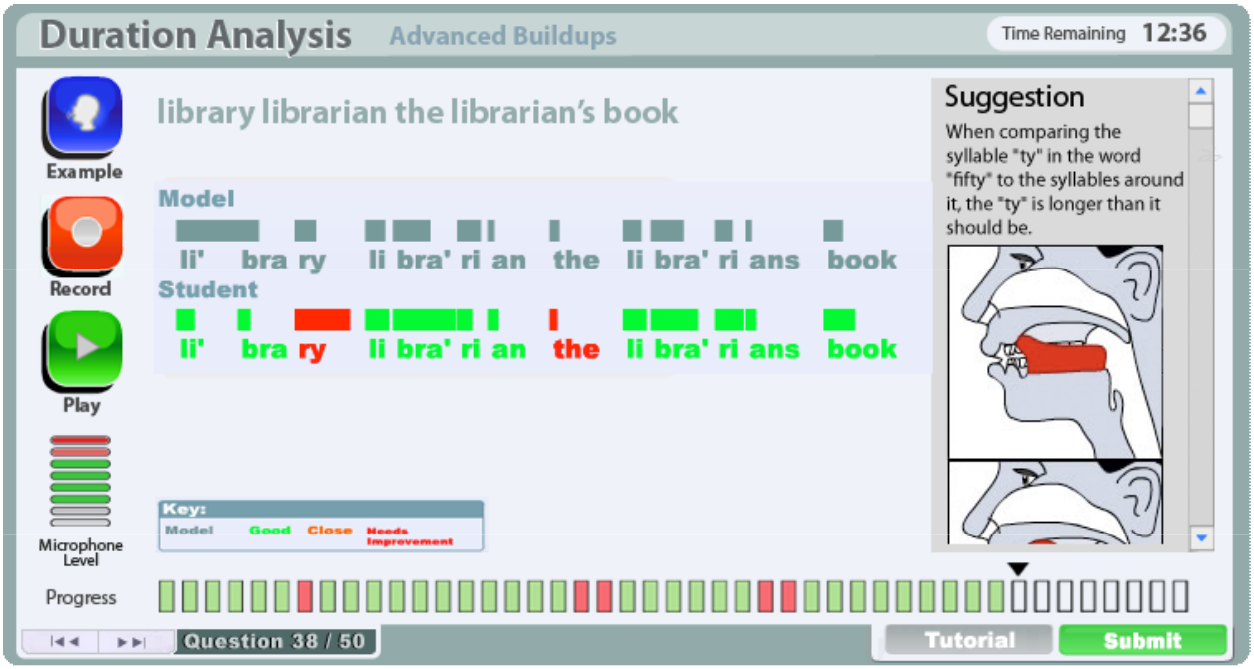


Figure 2. Proposed redesign. Errors appear in a dedicated feedback area on the right.

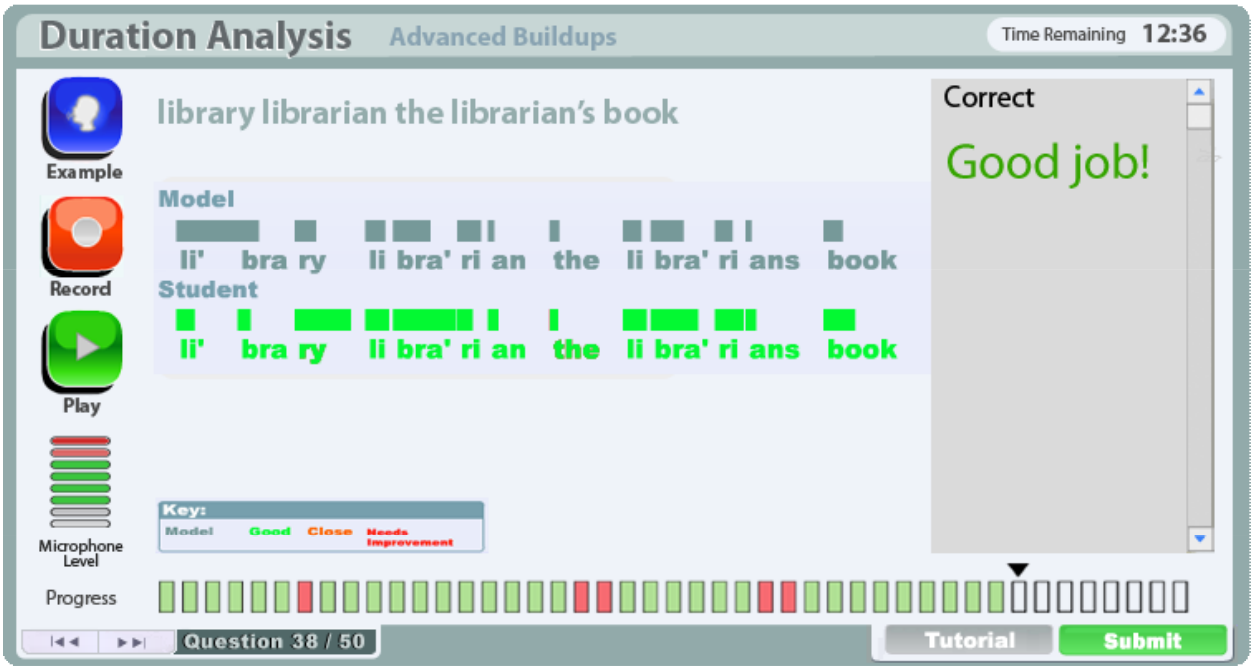


Figure 3. Proposed redesign. After a correct answer, a large "Good job" message appears to give a quick and easy to perceive indication that the response is correct and the system is moving to the next question.

Conclusion

Our team has discovered a number of usability issues with NativeAccent over the course of this project. We have outlined these issues at a high level in this report and included detailed issues in the appendices. We have also provided recommended solutions as well as redesigned screenshots showing how the current interface could be simplified.

Appendix A: Summary of Usability Aspect Reports

Below is a listing of all the Usability Aspect Reports (UARs) contained in this document. There are three tables, one for UARs from each of the following methods: Heuristic Evaluation, Cognitive Walkthrough, and Think-Aloud Usability Study. For each UAR, we list its number, whether it is a problem or a good aspect, the name of the UAR and a rating. Descriptions of the ratings are given below. The reports themselves appear the following appendices.

Severity Rating Codes
1 – Cosmetic problem only
2 – Minor usability problem (fix with low priority)
3 – Major usability problem (fix with high priority)
4 – Usability catastrophe (imperative to fix before release)

Benefit Rating Codes
1 – Cosmetic benefit only
2 – Minor good aspect (maintain with low priority)
3 – Major good aspect (maintain with high priority)
4 – Usability gem (imperative to maintain this aspect)

Heuristic Evaluations

UAR #	Problem / Good Aspect	Issue	Severity / Benefit Rating
HE-01	Problem	Confusing numbers on grades page	4
HE-02	Problem	No help or documentation	4
HE-03	Problem	Pitch analysis is difficult to understand and comprehend.	3
HE-04	Problem	Inconsistent Availability of “Next Lesson” Button	3
HE-05	Problem	Minimal indication of good performance	3
HE-06	Problem	Jump menu back/forward look like browser buttons	3
HE-07	Problem	Grade report is too cluttered	3
HE-08	Good Aspect	Questions can be skipped	3
HE-09	Problem	Ambiguous way to return to the course home page	3
HE-10	Problem	User action unclear on Start Screen	3
HE-11	Problem	Activity Report Text Does Not Speak User’s Language	3

HE-12	Problem	Intelligent Tutor name is unclear	3
HE-13	Problem	System does not speak user's language	3
HE-14	Problem	Many sections of the system aren't labeled in the users' language	3
HE-15	Problem	Breadcrumb labels are not clear	3

Cognitive Walkthrough

UAR #	Problem / Good Aspect	Issue	Severity / Benefit Rating
CW-01	Problem	NativeAccent course listing page does not provide sufficient cues for viewing a course	2
CW-02	Problem	Automatic advancement makes positive feedback perceptually invisible	3
CW-03	Problem	"NA-CMU" link label is not sufficiently descriptive or obviously denoted as the appropriate action for getting to the class home page	3
CW-04	Problem	NativeAccent does not provide sufficient cues for viewing a grades after submission	2
CW-05	Problem	"Jump To" Menu Does Not Look Like Standard Navigation	3

Think-Aloud Usability Study

UAR #	Problem / Good Aspect	Issue	Severity / Benefit Rating
TA-01	Problem	Pitch link is difficult to located on course home page	2
TA-02	Problem	User forgets to hit Record button	3
TA-03	Problem	Play button doesn't work when system auto-advances on a correct answer	2
TA-04	Problem	User confused by auto-advance	3
TA-05	Problem	Confusion on Suggestion box	3
TA-06	Problem	"Server Error" popup is confusing and does not provide sufficient help	2
TA-07	Good Aspect	User likes having a model speaker	3
TA-08	Problem	User skips question while waiting for feedback	3
TA-09	Problem	User has to search for "Submit" button at end of Activity	2
TA-10	Problem	User confused by Jump To	2
TA-11	Problem	Model speaker does not sound realistic to the user	1
TA-12	Problem	User unsure how to respond to "Speech does not match" error message	2
TA-13	Problem	User unsure how to interpret stress in Pitch Analysis	3
TA-14	Problem	User unsure how to interpret grades page	4
TA-15	Problem	User does not notice the model speaker for some time	2

Appendix B: Heuristic Evaluation Usability Aspect Reports

Below are reports on usability generated from a Heuristic Evaluation that our group performed. Each team member went through NativeAccent looking for issues that conflicted with a standard list of usability heuristics. We then compiled these issues and wrote up the most important findings.

Appendix C: Cognitive Walkthrough Usability Aspect Reports

Below are reports on usability generated from a Cognitive Walkthrough that our group performed on the NativeAccent site. This involved our team going step by step through the tasks of logging in, choosing a new activity, answering a question, and looking at the grades page. At each step of a task we looked for places where a new user would likely encounter an issue.

Appendix D: Think-Aloud Study Usability Aspect Reports

Below are reports on usability generated from a Think-Aloud usability study conducted with a new user. The user was asked to complete a pitch activity, a duration activity, and to view their grades. During each task the user was asked to "think-aloud" so that we could get a better understanding of their thought process.

Appendix E: Detailed Usability Issues from Contextual Inquiries

Below is a list of issues that were encountered during our contextual inquiry interviews. The list contains issues consolidated across all three users.

We plan to address these issues in our ongoing work.

General Issues

- Users had trouble understanding IPA symbols (e.g. course homepage, grades page, etc.)
- Users would click on large 'Carnegie Speech' logo on the NativeAccent start screen in an attempt to access the course
- One user clicked on the 'yes' button in the enrollment page without even reading the text
- Some users wanted to have the ability to practice first before 'being assessed' on their ability

Time-related Issues

- Users ran out of time before finishing the assessment, but were not aware that there was a timer
- Some users were unclear about the status of their completion – "Am I done because I'm out of time, or because I finished the last question?"
- Users were confused by the 'Jump To' button – failed to see the correspondence between course home page and order of menu items; also confused by left and right arrows (often misconstrued as having the same functionality as back and forward buttons in the browser)
- Users were confused as to where they were on the site
- Section Links on the course homepage (1 2 3 4) are ambiguous and redundant

Navigation Issues

- Users had difficulty getting back to the course home page from the Flash activities and other modules
- Users couldn't figure out the breadcrumb labels (NA-CMU; CS-Web)
- Users had problems recognizing blue hyperlinks displayed on blue background (e.g. NativeAccent start screen; Intelligent Tutor)

Intelligent Tutor

- Most users were confused about the purpose of the Intelligent Tutor
- One user thought it was a tutorial rather than a tutor
- No one seemed to treat it as a suggested lesson plan

- Users often did not recognize that the "Run the intelligent tutor to create a new path" link was clickable; also, they were confused by the meaning of "create a new path"

Grades Page (with Charts)

- Users had difficulty interpreting the two numbers underneath each bar
- Some users who read the heading "(expected probability of correct pronunciation…)" were confused that the probabilities went beyond 100
- Some users were looking for some sort of benchmark score for comparison

Activity Report (per lesson)

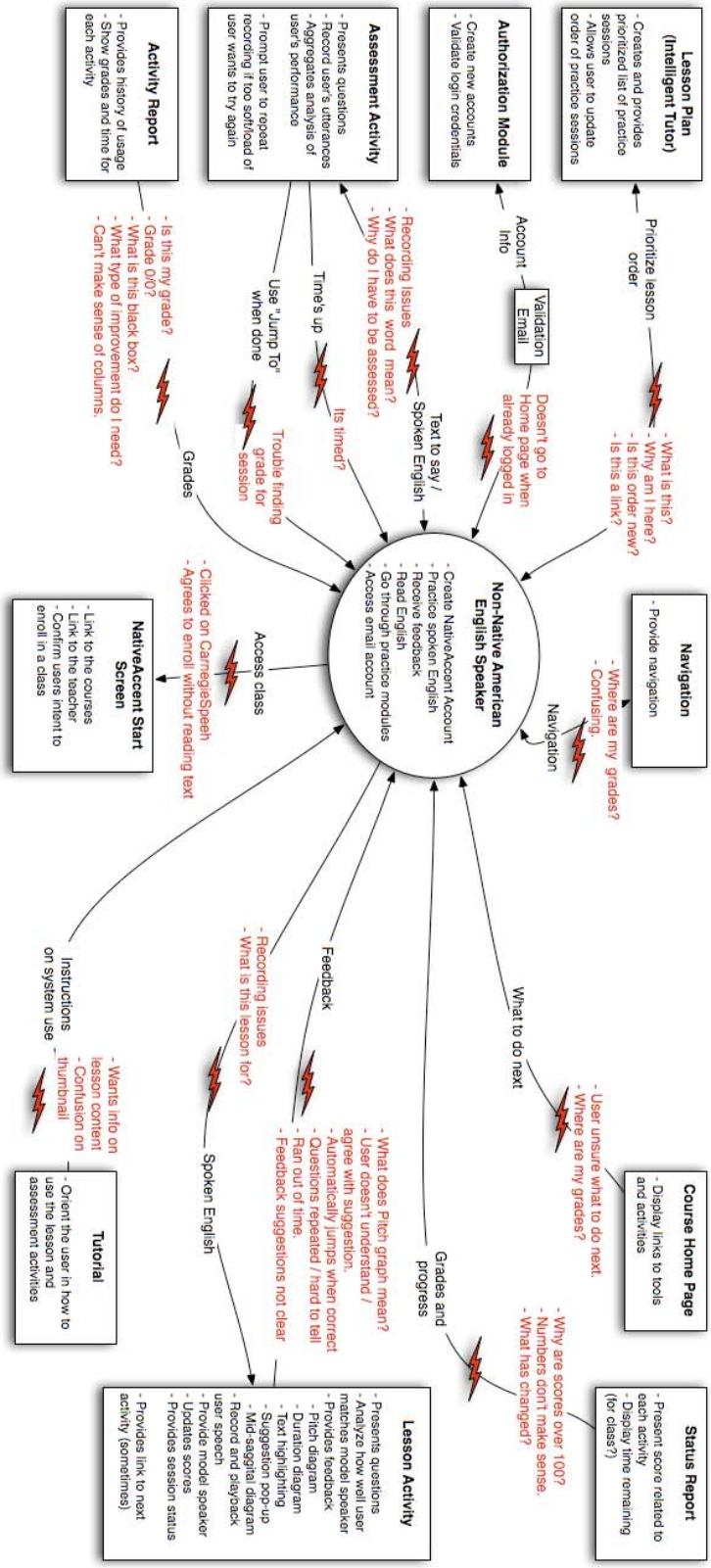
- Users were confused that "Your grade for this activity is 0 out of 0"
- Users saw a black bar for the table headers (had hidden text in it – Test Taken On; Responses) – misconstrued by some as a progress bar
- Users were looking for some sort of qualitative feedback, but numbers can't tell them what they need to improve on
- Some users completely didn't understand the Activity Report page

Lesson Activities

- Users who didn't first listen to (or chose not to listen to) the model speaker would generally be confused about what to speak
- Users were confused as to the purpose of a lesson; subtitles such as "Increasing Length Changes Stress Placement" (within Duration Activity) were not clearly understood
- One user encountered an error message: "Feedback could not be parsed due to an error in formatting. Please try again."
- Some users (who did not go through the tutorial) did not click on the 'record' button but instead hovered over it
- Some users questioned whether they were required to complete all 30 assessment questions
- Some users thought the model speaker was unrealistic; one user even *corrected* the model speaker for the word 'might' (model speaker audio was truncated at the end)
- Sometimes the 'Next' button () would take users to previously completed questions
- Some users would give up on certain questions that they couldn't perfect and were taking too long (e.g. 'monument monumental it's monumental size'; 'the libraries are going to lend us some books'), skipping to the next question
- The feedback in the Duration Activity overlaps other interface buttons
- One user encountered a "Moodle Service Error" message
- Recording function stalls if a user records for less than a second
- Tutorial: one user clicked on the thumbnails, thinking that they were interactive
- Some users were not clear about the distinction between tutor/tutorial
- Users have a difficult time interpreting the suggestion pop-up messages
- Some users wanted to do their own analysis of a successfully completed question (i.e. listening to their own recording and comparing it to the model speaker)
- Some users were confused by the graphical feedback and ignored it (one particular user tended to be a more auditory learner)

Appendix F: Flow Model

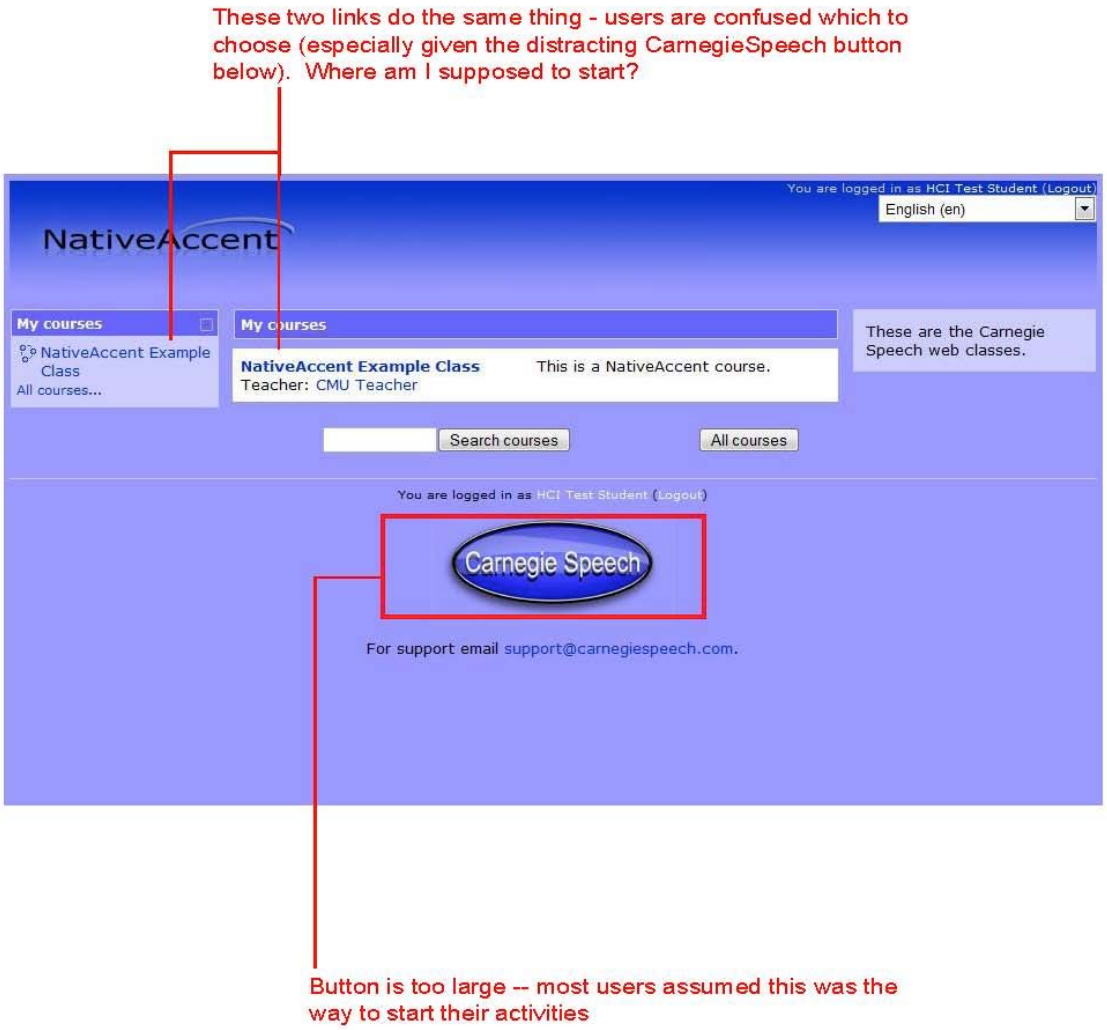
On the following page is a consolidated flow model that shows the users' interactions with various parts of the system. Red text and lightning bolts indicate problems where the user was confused or had difficulty using the system. The large boxes in the model represent abstracted components of the system, while the bulleted list inside a box describes its responsibilities. Lines between the user and large boxes represent communication flow. This model incorporates behavior from all three users; however, it does not represent any particular user.


Lesson Plan (Intelligent Tutor)
- Creates and provides prioritized list of practice sessions
- Allows user to update order of practice sessions
Prioritize lesson order
- What is this?
- Why am I here?
- Is this order new?
- Is this a link?
Navigation
- Provide navigation
Navigation
- Where are my grades?
- Confusing.
Course Home Page
- Display links to tools and activities
What to do next
- User unsure what to do next.
- Where are my grades?
Status Report
- Present score related to each activity
- Display time remaining (for class?)
Grades and progress
- Why are scores over 100?
- Numbers don't make sense.
- What has changed?
Authorization Module
- Create new accounts
- Validate login credentials
Account Info
Validation Email
Doesn't go to Home page when already logged in
Assessment Activity
- Presents questions
- Record user's utterances
- Aggregates analysis of user's performance
Time's up
Its timed?
Text to say / Spoken English
- Recording issues
- What does this word mean?
- Why do I have to be assessed?
Feedback
- What does Pitch graph mean?
- User doesn't understand / agree with suggestion.
- Automatically jumps when correct
- Questions repeated / hard to tell
- Ran out of time.
- Feedback suggestions not clear
Lesson Activity
- Presents questions
- Analyze how well user matches model speaker
- Provides feedback
- Pitch diagram
- Duration diagram
- Text highlighting
- Suggestion pop-up
- Mid-saggital diagram
- Record and playback user speech
- Provide model speaker
- Updates scores
- Provides session status
- Provides link to next activity (sometimes)
Activity Report
- Provides history of usage
- Show grades and time for each activity
Use "Jump To" when done
Trouble finding grade for session
Grades
- Is this my grade?
- Grade 0/0?
- What is this black box?
- What type of improvement do I need?
- Can't make sense of columns.
Non-Native American English Speaker
- Create NativeAccent Account
- Practice spoken English
- Receive feedback
- Read English
- Go through practice modules
- Access email account
Access class
- Clicked on CarnegieSpeech
- Agrees to enroll without reading text
NativeAccent Start Screen
- Link to the courses
- Link to the teacher
- Confirm users intent to enroll in a class
Instructions on system use
- Recording issues
- What is this lesson for?
Spoken English
- Wants info on lesson content
- Confusion on thumbnail
Tutorial
- Orient the user in how to use the lesson and assessment activities

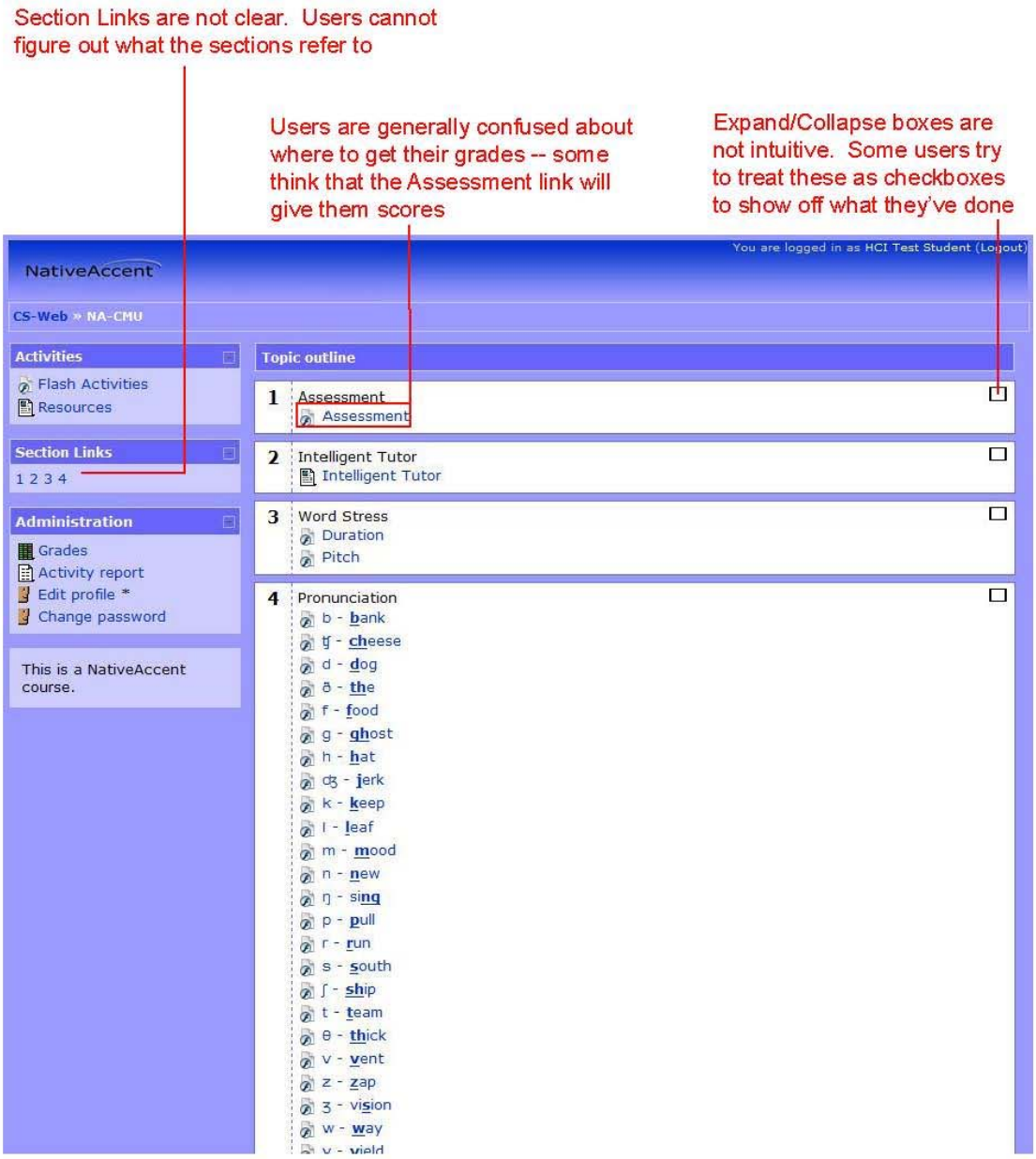
Appendix G: Artifact Models

The following pages show various screenshots of the system ("Artifacts") annotated with issues that users encountered on specific parts of the interface. The images and annotations reflect the state of the system at the time this report was compiled. Annotations articulate the consolidated issues experienced across all three users.

I. Home Page

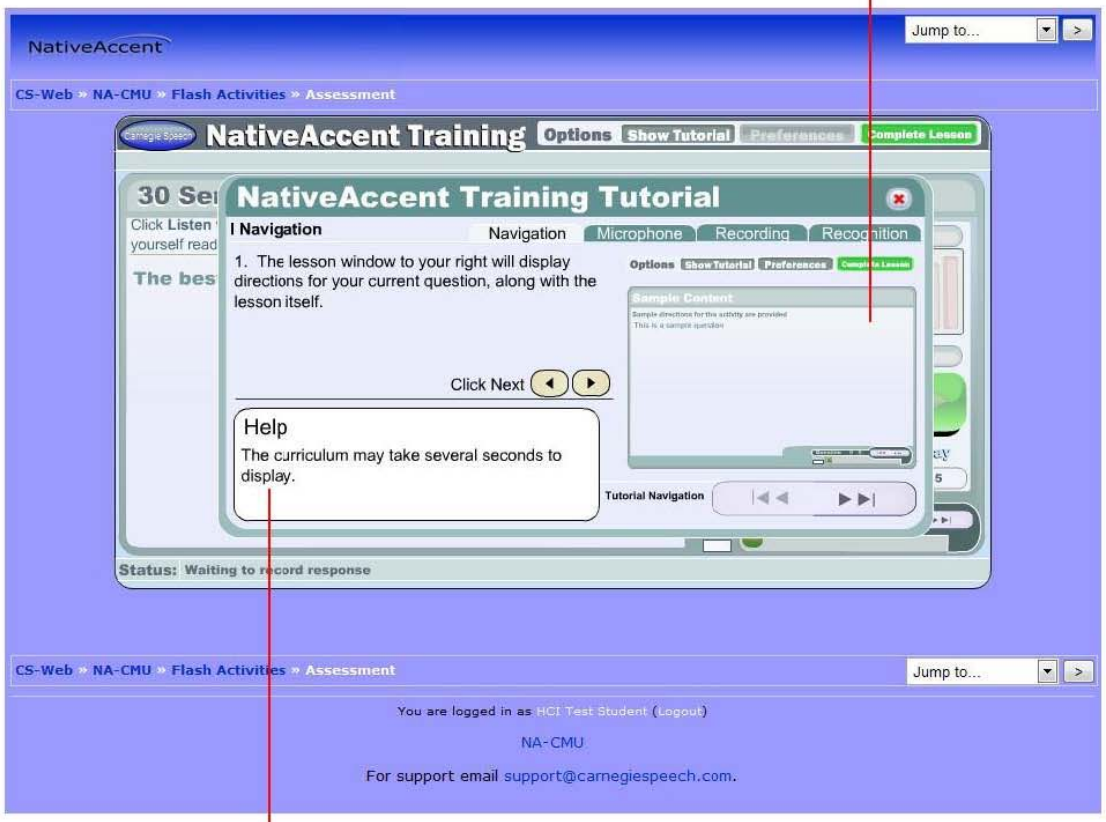

These two links do the same thing - users are confused which to choose (especially given the distracting CarnegieSpeech button below). Where am I supposed to start?
You are logged in as HCI Test Student (Logout)
English (en)
NativeAccent
My courses
NativeAccent Example Class
All courses...
My courses
NativeAccent Example Class
This is a NativeAccent course.
Teacher: CMU Teacher
These are the Carnegie Speech web classes.
Search courses
All courses
You are logged in as HCI Test Student (Logout)
Carnegie Speech
For support email support@carnegiespeech.com.
Button is too large -- most users assumed this was the way to start their activities

II. Course Page



III. Training Tutorial

Users assume the thumbnails are interactive and spend a reasonable amount of time trying to click on them only to realize that they are simply static images

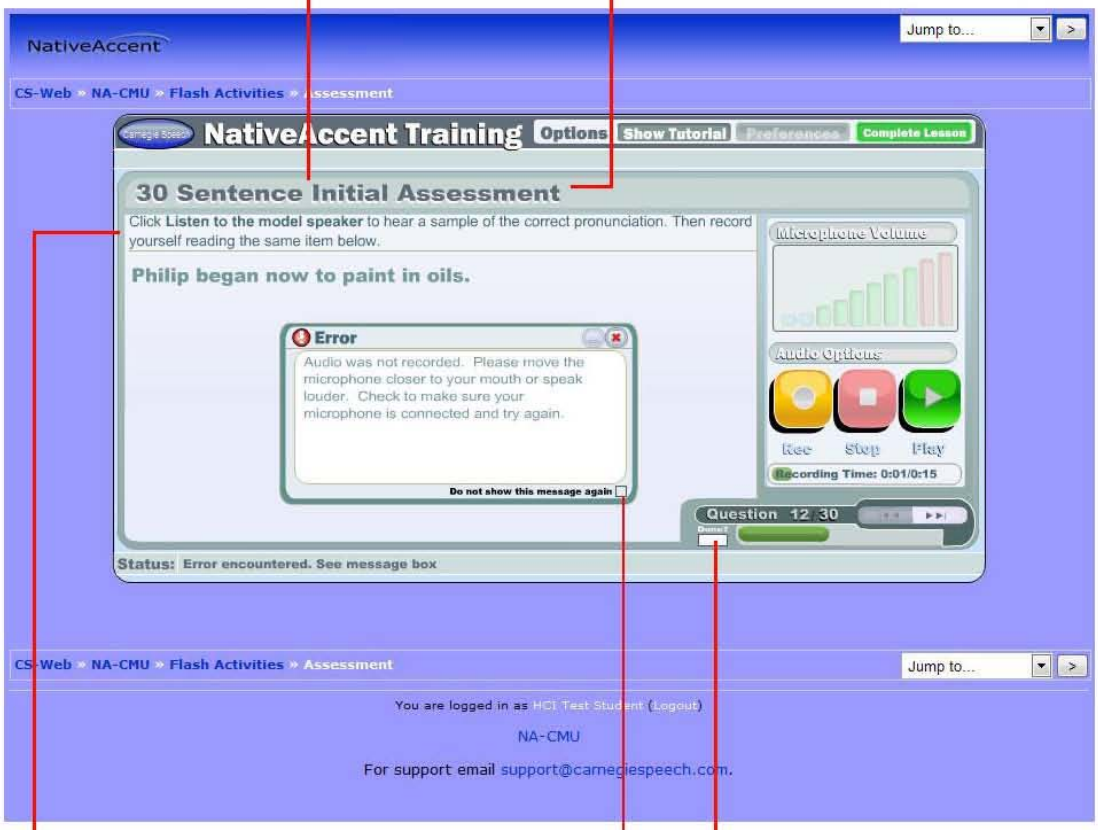


Users are confused about what the difference is between the tutorial and the tutor -- some expect the tutorial to tell them what they are going to be tested and graded on

IV. 30 Sentence Assessment

Users often ignore this text because it does not grab their attention (blends in with the background)

Some users aren't aware that an assessment is like a test -- they'd rather learn first



Most users don't feel that the instructions are detailed enough. Why am I here and what do I do?

This box appears to be an interactive check box that users can click on -- they become confused when they don't see any changes

Check box does not check -- users are often annoyed by the constant pop-ups

V. Training Activities (Duration, Pitch, etc)

User assumes the system is training them on something important, but they personally don't understand the value

Most users treat these buttons like the back and forward buttons in their browser -- they generally don't understand thow to use the 'Jump to...' menu (How do I get back to the home page?)

Users find the graphical feedback helpful

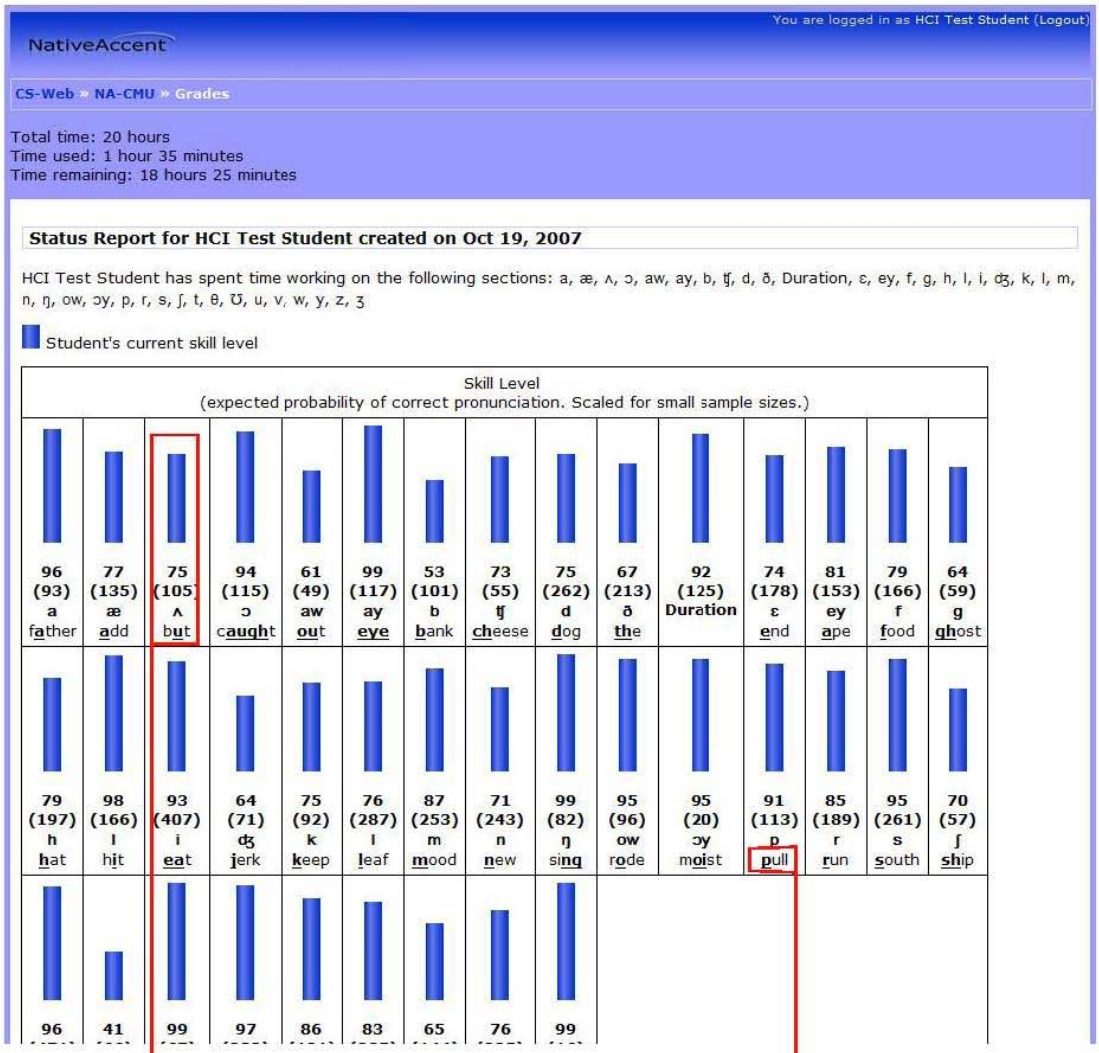
Users are not aware that the activities are timed until they actually run out of time

First time users often do not realize that they have to click the record button

Visual feedback is often not intuitive to some users -- most users listen to their response and then compare it to the model speaker

Some users aren't aware that the system automatically goes to the next question -- some want to hear themselves to make sure they actually did it right. Also, the back button does not consistently work

VI. Overall Grade Report



Users are unclear about these grades. Is this 75%? Why is my number in parentheses (105) over 100? How is that possible? How are the numbers in parentheses different?

Users expect these to be links so they can go straight to the lesson and improve their score

VII. Activity Grade Report