REQUIREMENTS SPECIFICATION

Kenzie Norris

Krystian Bednarz

Sam Asher

Preston Lee

Version 1.0

Sponsor: Dr. Okim Kang   Mentor: Daniel Kramer

Organizer: Michael Leverington

By signing this document the Requirements are expected to be the baseline for the following project

X _____   _____
    Dr. Okim Kang   Date

X _____   _____
    Kevin Hirschi   Date

X _____   _____
    Michael Leverington   Date

X _____   _____
    Kenzie Norris   Date

Table of Contents

1. Introduction

The United States is the #1 study abroad country in the entire world, where people around the world seek to better their education through the schools and universities held there. Considering that English is the main language in the United States, the main hurdle people must overcome when achieving an education in the US is the understanding of the English language. Learning a language is much easier if taught from a young age, but individuals trying to learn a new language later on is much more of a challenge, as language learning becomes more difficult with age. Certain applications can help with vocabulary and grammar, Duolingo for example, but key factors in speech like contextualization aren't quite realized through such existing applications. Considering that the industry in language learning as a whole is so large that it must encompass everyone pursuing literacy, a better functioning language learning alternative would serve the general public on a large scale.

Our project will be a website application that consists of various learning platforms that aim to teach contextual awareness, as well as a mastery of pronunciation with words in the English language as a whole. To facilitate this learning process, our website will have multiple gaming elements with applications geared towards language learning. Expressing the various stylization of certain words, and language library with a search feature for specific word contextualization, are both features used to aid users in understanding which word is used in specific circumstances. Open world gaming elements are present in our project scope as well, where interactable objects can be referenced and stored in a user's own personal library for future reference.

Our clients are Okim Kang and Kevin Hirshi, both members of the NAU department of English. We are one of the many capstone teams that worked to pursue language learning over the years. The focus of our particular capstone project is to aid transferring students to NAU learn how to speak the English language in the most efficient and fast process possible. Mastery of pronunciation and word contextualization would also be included in this process, where incoming students can both better speak the language and understand the language being told to them, maximizing their education at NAU. We intend to take the years of research our clients have made and use it to develop a gamified language learning tool, that teaches both the fundamentals of the English language, and teaches the more nuanced intricacies of the language overall.

To that end, we must highlight what exactly must be present within our programming project. What systems must be put in place, what features must the project have, and what the minimum viable product should be. Each component must be analyzed and understood as a necessity to achieve our client's goals for the project, and how we go about delivering results must be explained as well.

2. Problem Statement

Our sponsors are deeply involved in the language department here at NAU. They are both fascinated by linguistics, specifically in how we acquire new languages. In their work they recognized two gaps within the language learning process that they would like to try to fill with this product. They believe that when learning a new language it is crucial to not only keep the user engaged in the learning experience, but also to allow the user to interact with the language in natural ways that enhance their retention and provide them with a deeper understanding of how the language is spoken practically in day to day life.

One of the most common failures of students who are learning a new language is a distinct lack of engagement. Learning an entirely new vocabulary, a new set of grammar rules, new sentence structures, as well as all the small intricacies of a different dialect is not only overwhelming, but also can be a very dry and frustrating process. This is unfortunate because as many people who have succeeded in becoming proficient at another language will tell you, consistency is key. Learners who are able to build a strong routine into their language learning journey are the ones who see the most success. Our sponsors recognize this and want to make an attempt to fix it. They believe that learning a new language can be fun and invigorating if presented in the correct way, and this in turn is what will keep engagement strong. If the process itself is rewarding then the students will be much more likely to return and continue their education. For this reason they would like to introduce a gamified aspect to the product, which gives the program some fun elements that the user can enjoy and continually improve with.

Another area that almost all language students struggle with is pronunciation. Being able to read and write words and sentences in another language are important skills, but being able to fluently speak and pronounce words is what separates an average student from a great one. When you're unfamiliar with a language, listening to a native speaker talk will often have you wishing that they would slow down, and it is not always a given that their enunciation will be perfect. Still, to other native speakers they can be understood. Many

languages have small spoken subtleties that change how words are said in different contexts, and this is not an easy thing to teach. There is also a natural flow to most languages that is simply impossible to understand without immersing yourself in it, and hearing sentences being spoken in a natural setting. No two languages share exactly the same sounds or ways of pronouncing them, and inexperienced learners may find this to be a large hurdle to overcome. This is the second key problem that our sponsors hope to overcome with this product.

3. Solution Vision

Our solution to these language learning problems is a gamified language data explorer that helps with pronunciation through analysis of data, or PANDA for short. It is a browser based web application that is connected to a large audio database provided to us by our sponsors. The database is a large library containing segments of speech from native speakers of a language, taken from a variety of sources. Our application will need to contain a robust and efficient search feature capable of accurately retrieving relevant results from the database based on the user's entry and displaying them in a user friendly and organized way. This is a key feature of our application, and will provide a lot of utility to all users.

This system will initially be used mostly for research purposes, and will provide a useful tool for allowing our sponsors to further test their language learning theories. This will be a strong foundation for an application that they have the opportunity to add to and improve upon in the future, as they come up with new features that they feel would benefit users. This software has the potential to change the way that we look at learning languages. It could become a powerful user-friendly tool used by language learners and academics alike to easily and methodically explore a language, and test themselves on their applications of it. We hope that our project not only helps people in their path to knowledge, but gives them fun and effective ways to gauge their knowledge while doing it.

The following are the main features that we envision in our final product, further details about each are explained in the requirements section following this one. Figure 1 shows how each of these components interact with each other.

3.1: Library Search Function

- One of the key requirements of our project. Allows users to explore the accompanying database and interact with the results.

3.2: Administrator Page

- This will be a useful page for our sponsors, or whoever they designate as administrators to control and edit different aspects of the application. It will include functionalities that allow the administrators to interact with the database and with user accounts and groups.

3.3: User Profiles

- Contains user specific data, such as learning statistics and progress, as well as points. Also where the user will be able to access their personal notebook.

3.4: Practice Page

Students will be able to select a word from the results page to reach the practice page OR go to the practice page manually and the game page will select words to work on automatically from the user's personal notebook. The practice page will contain the three following games to help users work on difficult words:

- **The Pronunciation Game**
  - A game that allows users to practice their pronunciation of words by listening to and repeating a word. The user will be able to rate themselves to help track their progression.
- **The Context Game**
  - A game that challenges users ability to use context clues to determine the accurate word for a given sentence. Allows the user to earn points, and to see the correct answer if they are wrong.
- **The Dictionary Game**
  - A game that also works on a user's contextual reasoning skills. Focuses on determining definitions for words based on their context in a sentence. Users are also awarded points for this game.

3.5: Leaderboard

- A good way for users to use friendly competition, either between all players, or a select group, in order to drive their progress.

Figure 1: Shows the relationships between each of the key components of our gamified data explorer application.

- Each user's own avatar icon will be displayed alongside their performance statistics based on various language learning categories. It is important to note that no indication of user's personal information will be displayed on the leaderboard for privacy, only the user avatar in addition to their performance stats will be shown. Full names and usernames will be omitted from the leaderboard, as examples.

4. Requirements

For the Domain Level section, we will detail the requirements for the system described in the solution section, as well as expound upon it through functional, performance-based, and environmental requirement sections.

4.1: Functional Requirements

Functional Requirements are ones that are measurably there or not. Having them laid out in this form allows us as a team to know what must be functionally working by the end of the project. To make this easier to digest, we made a figure showing the 6 major categories covering what'll be expected of us by the end of the project.
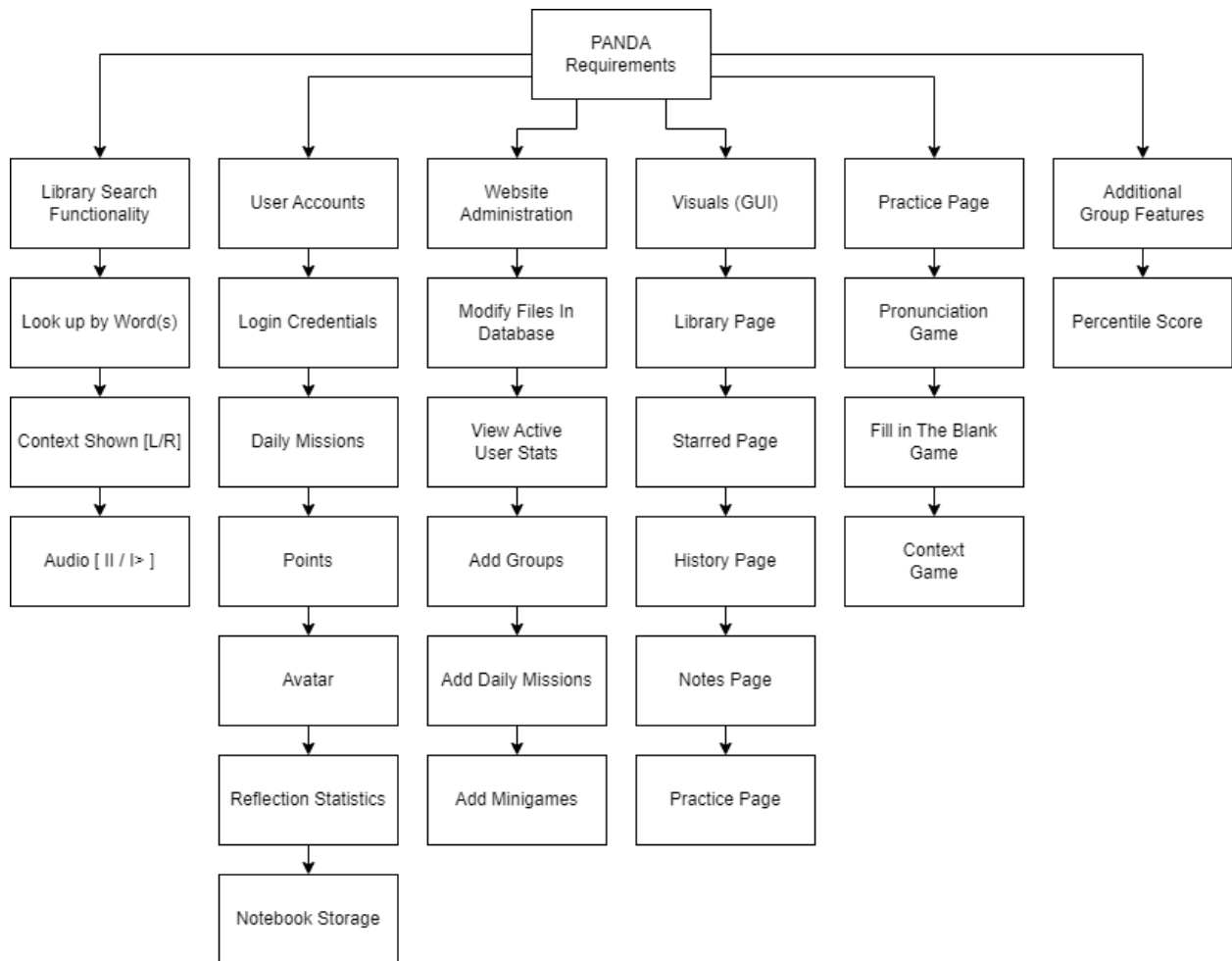
Figure 2: The following figure shares each category at the top measured left to right, top to bottom in order of steps for completion, as well as fundamental importance and dependency.

To better detail these steps, the following section aims to expand with specifics to better guarantee requirements are met.

4.1.1: Library Search Functionality

As the website will be built for aspiring English learners, it's important to provide them with the ability to search through our database to find what they need. The contents should be displayed to their liking, as well as interactable to provide full context for easier application.

- The system should be able to search for a singular word, or a pair of words, providing each entry that satisfies the query, without duplicates.
- The user should have the ability to hone in on select words to see a number of words in either direction of the query. This is to make the user able to hear the example sentence closer to the moment the word is said, rather than in its entirety every time.
- The system should have the ability to start and pause the audio segment to prevent annoyance, as well as train of thought derailment on the user's end.

4.1.2: User Accounts

Wanting the user to feel inclined to come back, the website must include an optional saving feature through accounts to allow for a more personal, educational, and enjoyable experience. The account must:

- Have a username and password. The password must not be stored without some form of encryption to prevent damage from data leaks.
  - Password resetting will be done through administrators, where a password reset request can be sent from a user to administrators.
  - Authorization of password reset for specific users will be at the discretion of the admin responsible for the password reset.
- Store details for completion status in daily missions.
- Store the points earned as to incentivize the user.
- Store which avatar is being displayed to enhance user customization and uniqueness.
- Store some amount of statistics to help the user reflect and figure out where to improve.
- Store any starred (saved) sentences, a brief history of interaction with sentences, as well as notes written by the user.

4.1.3: Website Administration

With the need for adding content and managing what's available, an administrator role in its infancy is at the very least required for the project's end. The administrator should be able to add, remove, and edit anything using the GUI, allowing it to make direct changes in the database upon changes.

- As said, the administrator must have the ability to make changes in the database to make it to their needs.
- The administrator needs access to active user stats to both see how the site is being used, as well as how the community is enjoying it. This is required to figure out where to improve among other things.
- As the client wants groups to be something controlled, the ability to create one for testing purposes needs to be only accessible via the administrator.
- With daily missions needing to be uniquely challenging, the ability to add, edit, and remove these missions would be critical to increasing the longevity of the website.
- Wanting the ability to add onto the website, new games should be addable via some means.

4.1.4: Visuals (GUI)

Being a gaming interpretation of the learning model, the visuals to show off these capabilities must exhibit these properties. Sectioning each part in a manner that helps the user better understand and navigate the environment is crucial to the longevity of the website.

- Wanting the user to be able to find new words they've heard of or explore other similar sentences with a word they've seen, the library page is where the user is encouraged to explore.
- With the user being bound to like certain sentence examples more than others, not to mention wanting to save examples to come back to later, the star feature is meant to provide this saving need.
- As many users are bound to forget something they just saw, having a history feature allows them to quickly get back on their train of thought and speed up the learning process.
- Having a notes page, while simple in theory, allows users to express their deep thoughts in the same application they learn in. This allows for faster learning,

compiling a personal dictionary of sorts, while doubling as another reason to come back.

- The website would simply be an exaggerated dictionary without the gaming aspect, so having the practice page allows for a quick and easy experience with the words in question while having fun in the process.

4.1.5: Practice Page

With the practice of using the words being the main attraction of the website, they need to be numerous, and more than just flashcards. To do this the team came up with a couple simple games to start testing the user's ability through pronunciation, filling in the blank, as well as context definitions.

- The pronunciation game must provide the user with a word in a given context, ask for a non-saved recording with multiple retries of the user saying said prompt, and then provide the actual recording to allow the user self-reflection when comparing the two recordings. To aid the user in self-reflection, they are encouraged to rate themselves and allowed to retry as many times as they'd like.
- Filling in the blank follows the thought of knowing where a word fits, and thus knowing where a word feels out of place. To teach this, the user will be provided with 3-4 options where a word is missing. The audio can be played up until the blank, and the user will be rewarded with points based on getting it correct. Upon completion, correct or incorrect, the user will be able to hear the full sentence and be prompted to try again with a different example.
- Discerning a definition based on context clues is a very useful skill taught at a young age that even adults use day to day. To aid in this process, as words change with time, the game selects a word in a sentence given the context. Providing the user with possible dictionary sources online encourages the user to come up with their own definition and compare it with the dictionary. Upon completion, they'll have the opportunity to add it to their notes to make an extensive list of personal definitions that can be ingrained deeper than a simple flashcard that's easily accessible to that personal user.

4.1.6: Additional Group Features

To help with user engagement we believe that it's important to add features that will reward interaction, and give feedback to help learners track their progress. Giving them a

motivational percentile score to look at to compare with others would both provide a goal, a reason to come back, and gratification to see the number go up.

4.2: Performance Requirements

With the functional requirements out of the way, the performance requirements are going to provide a measurement in a pass/fail manner that could easily be testable to keep the clients wants in order.

- **Library Search Functionality**: The search happens in under 5 seconds before the first or all results return.
- **User Accounts**: The user account will have a maximum of 50 definitions and 50 starred sections.
- **Website Administration**: The changes to the database happen within 10 seconds of the request.
- **Visuals (GUI)**: The notebook allows for swapping interchangeably between pages within a reasonable time frame (5 seconds).
- **Practice Page**: The games feel fluid, minimizing page updates. The page only updates once to load each game.
- **Additional Group Features**: Percentile is quickly calculated, and summarized within 1 second to be lightweight.

4.3: Environmental Requirements

The key environmental requirement of our project is that we cannot use Gmail authentication for our password system as it needs to be accessible by students from China. The requirements listed in this section have been made from meeting with the client directly, and have been made to show what is needed in the final product at the minimum. The website will be run on HTML with CSS and javascript, using PHP and MySQL in the backend on an EC2 AWS instance.

5. Potential Risks

As seen in Figure 3, SQL injections, password encryption, and scalability are all three potential risks to assess before giving the project an all-clear.

Risk	Impact	Possibility
SQL Key Leak	High	Very Unlikely
Password Leak	Moderate	Likely
Lack of Scalability	High	Very Unlikely

Figure 3: Possible Risk Factors

5.1: SQL Key Leak

With the database being reliant on mySql and AWS, it is important to make sure these keys are not leaked as to make sure that the data stored there isn't erased, ruining a ton of user's progress. To prevent such threats, we must keep said keys secure in the right hands, as well as have a backup in case of catastrophic events.

5.2: Password Leak

Password security is vital to any application that is expected to handle user accounts. User passwords will need to be hashed and salted as a minimum safety requirement to ensure that any kind of personal user data does not end up in the wrong hands. We don't believe that malicious attacks are likely, but must ensure the safety of the data that users entrust to us. This is a potential vulnerability that we cannot ignore, and cannot afford to discover too late.

5.3: Lack of Scalability

Choosing AWS, we shouldn't need to worry about the inability to scale up. But if we were to find ourselves in this said predicament, we would need to update everything to match the new database. Of course, this is assuming the host was the problem. If we manage to not make the system modular enough, it would be too hard to add onto, and that would also be a potential risk. But as we already are planning for a modular design, this shouldn't be a likely outcome.

6. Project Plan

In terms of future tasks pertaining to project completion, we have determined a few key sections that must be completed. Certain sections are dependent on others, meaning prior sections must be completed beforehand before beginning work on future sections. Figure 4 is a Gantt Chart that displays such dependencies using arrows, where completion of one section allows for starting another.
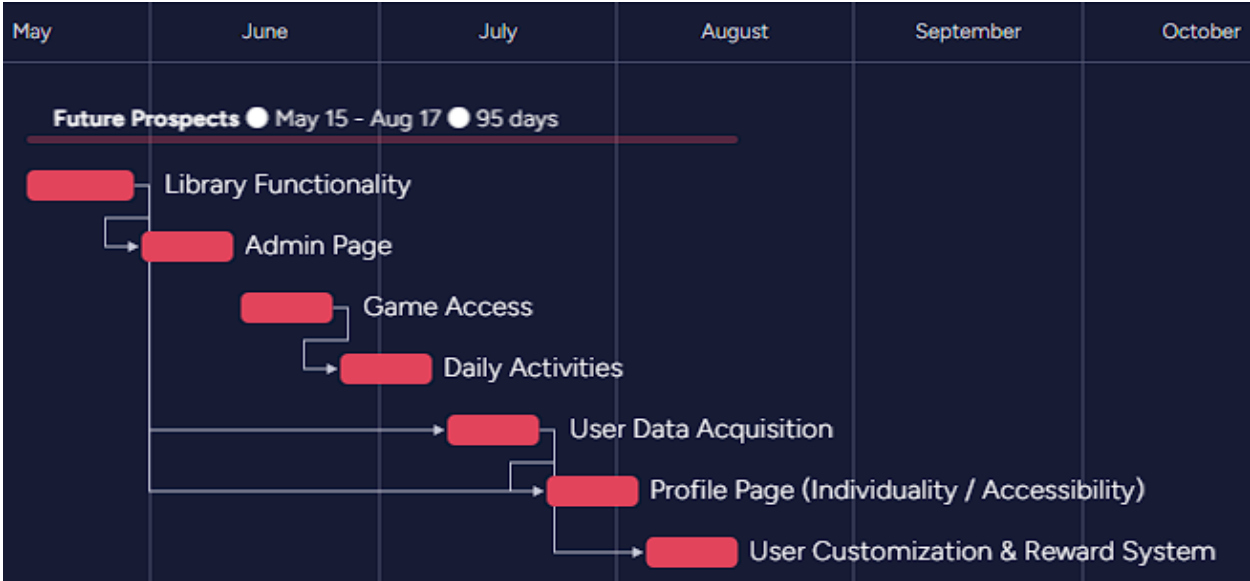

FIgure 4: Gantt Chart of Future Prospects

Each section is assigned two weeks of working time for now, where the sections that satisfy the more dependencies are placed earlier in the chart. This current rendition of the Gantt Chart is subject to change, depending on workload and complications that could arise in the future.

We intend to begin working on the library functionality first, doing so to ensure that all necessary data can be stored in the backend database for use throughout the project. Once the library is in place, work on the admin page can begin, where we can assign necessary permissions to both view and edit data within the database. Game access will consist of stub links to various possible gamified elements to be used initially, but further implementation of what games can be played will eventually be stored in a readily available page for users to access. Once in place, daily activities can be created and referenced based on the games we will have available. Before setting up the profile page

for regular users, we need a means to acquire data on what progress the players have made throughout their language learning journey. It is important to set up a means to capture all user progression first, as progression is intended to be displayed using graphics on a user's profile page. After the profile page and user data can be acquired, user customization can be worked on, as well as a reward system. This section is more of a stretch goal assuming the rest of the sections are mostly completed, but we would like to include a customizable player avatar to encourage language learning on our site.

7. Conclusion

The language learning tools present right now help with the language learning process, but do not fully encompass the intricacies of language as a whole. Basic vocabulary and grammar can be grasped, but the overall understanding of how to speak a language as a whole is not fully realized. Our team is tasked with the development of a gamified language learning tool to further the learning of languages beyond what is already in place, more so involving pronunciation and word contextualization. Through usage of gamified elements, language library searching features, and audio prompting, we are certain that our method of language learning will be much more effective.

Throughout this document, we have highlighted each component in both the concept of language learning, and what exactly our project needs to have to function how we intend it to function. The main problems with language learning and the solution we have to fix such problems are made known. The requirements for what features our project must have are identified, and will serve to determine what needs to be done by the delivering of our finished product. What possible risks we may encounter through project development are now known, highlighted here to determine what actions must be taken to prevent any possible negative outcomes throughout project development. Finally, we have developed a scheduling system to ensure key components are developed in a timely manner.

Our team intends to make language learning as effective as possible, while making that process more enjoyable. Learning a new language can be a difficult process, and is a requirement to further education in foreign lands. The pursuit of education should not be halted by not knowing how to understand what is being taught. Language learning should be accomplished through a seamless and effective way, this being the main goal of our project. We will make this process a reality to better the pursuit of education for everyone, opening the door for aspiring students everywhere.