

Building and Deploying a Paradigm Generator with the Apertium API

Matthew Parker
Linguistics 073

Project Scope and Goal:

Creation and deployment of an extensible, open-source verb paradigm generator using the Apertium translation API hosted at beta.apertium.org. The end goal of this project is to enable linguists to, for any verb in a supported language, gain an overview of the different ways in which that verb is used in a target language.

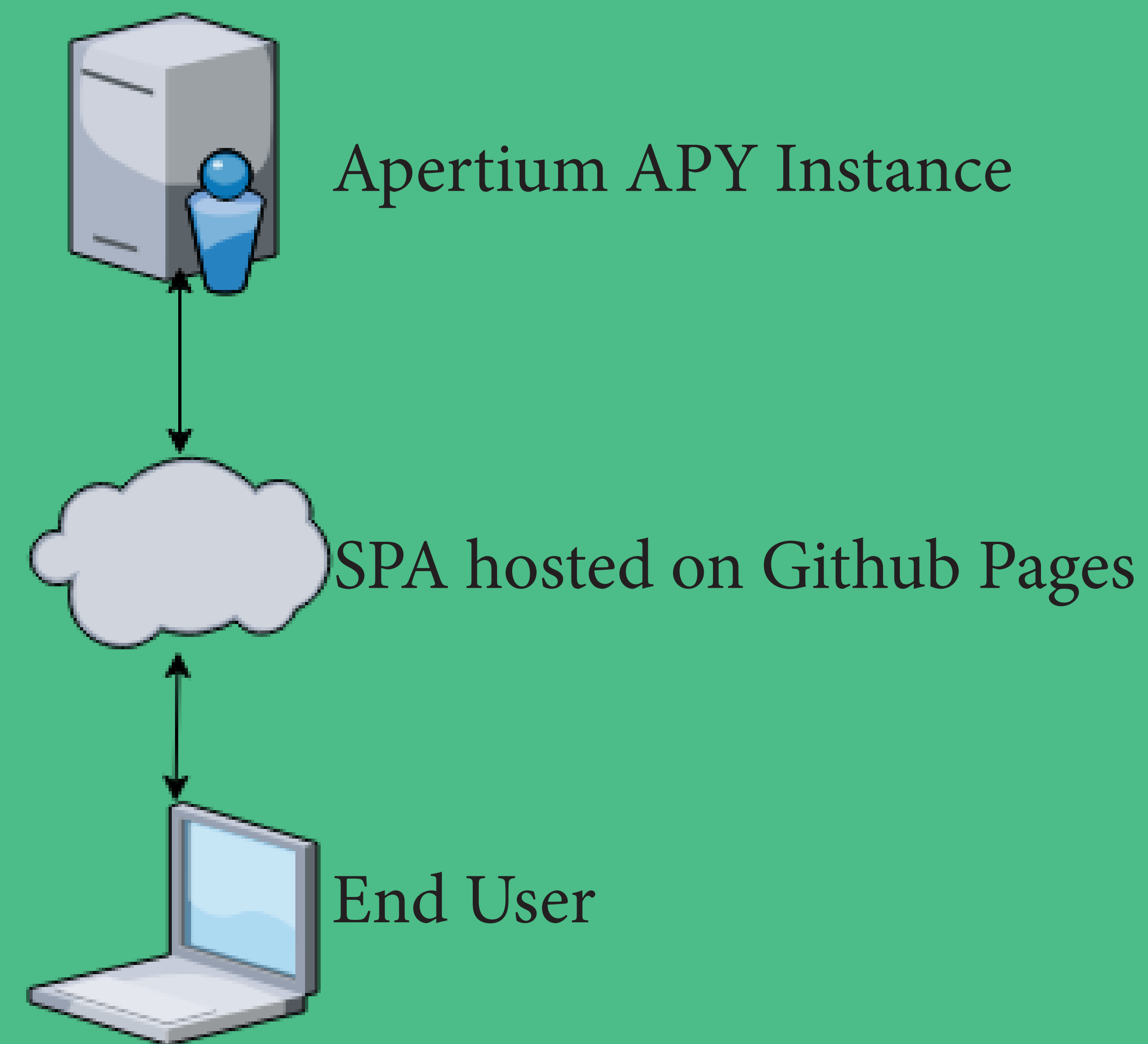
Methodology:

Use open-source framework React.js to create an SPA allowing an end user to enter a verb infinitive and receive different verb paradigms for the verb.

Use Cases

- Identifying unfamiliar forms of verbs seen in the course of linguistic research
- Providing an overview of the various forms of verb conjugation that exist in a language
- Helping to display patterns in verb conjugation – for instance, this could help identify if first-person verbs are conjugated the same way in past tenses, but differently in the present and future tenses
- Could be used to train black-box translation systems by being leveraged to create a corpora of verb conjugations.

Paradigm Generator Architecture



Project Fit

This project addresses the distinct lack of an adaptable, open-source verb paradigm generator. Verbix offers multi-language paradigm generators, while assorted projects offer open-source single-language generators, however, no projects fit both criteria. Ideally, this project will contribute a FOSS solution to the field of verb paradigm generators, and will be extensible to those working with undersourced languages. A FOSS solution to this problem should offer comparable performance to current closed-source solutions.

Extensibility and Improvements

- Separate out tag logic into a separate class to allow easy modification for different languages
- CSS/Styling/General prettification
- Add language selection feature
- Add support for other parts of speech – nouns first, and then adjectives and adverbs
- Move to a test-driven development lifecycle
- Host APy instance to minimize dependency on external infrastructure
- Make an external API available for developer use – Verbs as a Service

Evaluation

The planned evaluatory criteria is to be able to enter any verb in a test language (Spanish for now) and receive a list of verb paradigms. This goal is still very much a work in progress. Current issues include dealing with asynchronous API calls and ordering in the React framework. While not part of the planned criteria, a successful deployment of this project would involve maximizing accessibility to less experienced developers looking to contribute to a new language. Other evaluatory criteria include accuracy of translations and paradigms generated.