Computational Analysis of English Verb Inflection Using IPA and FSTs

Riya Singh and **Yara Hajeissa**

Abstract

We present a computational study of English verb morphology using Finite State Transducers (FSTs) and the International Phonetic Alphabet (IPA). By constructing FSTs representing the production of both regular and irregular verbs, we modeled the transformation of IPA-transcribed verbs into their past, progressive, present (third-person singular), and past participle forms. Additionally, we developed inverse FSTs to map inflected verbs back to their base forms with tense markers. Evaluation on a standardized dictionary of 269 American English verbs demonstrates that a hybrid words-and-rules FST achieves high accuracy across all tenses, outperforming purely rule-based models and a morphological analyzer. Our results provide a computational verification of Pinker's words-and-rules hypothesis and highlight the challenges of modeling irregularity, orthographic variation, and IPA standardization in English verb morphology.

1 Introduction

In linguistics, the International Phonetic Alphabet (IPA) is a standardized phonetic notation system created by the International Phonetic Association for academic use. The IPA uses a set of unique symbols that correspond to the distinct sounds found across human spoken languages (International Phonetic Association). This sound–symbol mapping provides a phonetic representation that is consistent across all languages. Within English, two major phonetic standards appear, American English and British English. In our work, we work with the American English conventions.

We rely on the IPA because it allows us to analyze phonological rules rather than orthographic patterns, enabling us to focus on the actual pronunciations of American English verbs. This is especially important because many verbs exhibit inflectional patterns that involve phonetic-not orthographic-changes. By examining the IPA forms of verbs, we aim to better understand the structural patterns underlying regular and irregular verb morphology. In particular, we are interested in how these patterns intersect with what we know about child language acquisition patterns. Young children are accurate with irregular forms very early on, as they often memorize irregular verbs before they have any productive rules to apply to regular verbs (Marcus et al., 1992). Accuracy then temporarily drops once they infer the general rules and begin over-regularizing (goed, taked) (Marcus et al., 1992). This makes irregular verbs an especially rich domain for studying how children balance memorization and rule formation.

Motivated by these linguistic and developmental questions, our research question is as follows: With what accuracy can we map IPA-transcribed verbs to their corresponding past, progressive, present (third-person singular), and past participle forms using a dedicated Finite State Transducer (FST) for each transformation? Additionally, with what accuracy can we predict the infinitive form and tense of inflected verbs?

2 Hypothesis

Based on the properties of English verb morphology, we expect the progressive tense to exhibit relatively high accuracy, as it follows a consistent set of rules and has few irregularities. Present tense verbs include some irregular forms, but these are less frequent than the irregularities found in the past and perfect tenses. Consequently, we anticipate lower accuracy for past and perfect forms, since handling irregular verbs requires hard-coded exceptions, and it is not feasible to account for every irregularity, leaving more room for errors.

3 Methodology

3.1 Regular Verb Inflection FST

To answer these questions, we implemented a Finite State Transducer. To construct our FSTs, we used the provided fsmutils.py and fst.py libraries we utilized in Homework 3. Initially, we only followed the regular patterns of each verbal tense. In the present tense we added /s/ after voiceless wordfinal consonants, /z/ after vowels and voiced wordfinal consonants, and /ɪz/ after wordfinal sibilants. In the participle and past tense, we added /t/ after voiceless wordfinal consonants, /d/ after wordfinal vowels and voiced wordfinal consonants, and /ɪd/ after wordfinal alveolar consonants /t/ and /d/. In the progressive tense, we simply added /ɪŋ/ to the end of every verb. It is important to note that we initially represented the orthographic doubling sometimes seen in verb inflections (eg. run → running) when the final syllable of a verb is stressed and follows a CVC construction. We succeeded in building a model for this but later abandoned it for the sake of standardization, which we will continue to discuss later. This phenomenon is purely orthographic and is not realized in the phonology of the English language. Initially, the results of our models were promising, but our model's performance began to drop off as we tested our model against more irregular verbs, since our basic model was not built to handle irregular constructions.

3.2 Irregular Verb Handling

From there, we took on a "words and rules" approach. We began to construct a separate FST that would deal with the irregular verbs of English. We located an extensive, but incomplete list of irregular verbs in their infinitive form, past tense, and as a participle. We transcribed each irregular form in the IPA to be used in our model. To implement the irregular model, we used a branching strategy such that the states of this FST would be organized alphabetically. This means that each initial state, spanning all the symbols of the IPA found in English, represented each initial phoneme of the verbs handled by this FST. These initial states would not have any output, and branch off into secondary states that represented the second phoneme that followed by the first one it stemmed from. This branching sequence would continue until the final end bracket is found. This is when the output is given, with the corrected irregular inflection in the corresponding tense. We combined these FSTs by traversing through the irregular model first, and if its output was blank, we would then traverse through the regular model.

3.3 Reverse Inflection and Morphological Analysis

We then constructed another FST going in the opposite direction. This FST takes in a verb in any inflected form and outputs its bare form with a tense marker attached to it (eg. #dɪvɛləpt# → #dɪvɛləp# PST/ PST.PTCP). Similarly, we implemented a "words and rules" model because the purely regular model cannot account account for irregular constructions such as #sæŋ# (sang) as the past tense of #sɪŋ# (sing). For the irregular model, we used a similar branching approach as we did in our forward direction models. The input was the irregular verb of any tense, and the output was the bare infinitive with the appropriate tense marker.

Both the verb inflector and the reverse inflector prompted the user to input a verb. English spelling was allowed as long as it was included in our dictionary, otherwise the user would provide their IPA transcription of that verb. The inflector would return the bare infinitive, past, progressive, present, and participle form of that verb. The reverse inflector would take in a verb and return its bare infinitive form and its tense marker (BARE, PST, PROG, PRES, PST.PTCP).

The regular model worked in a more patternized fashion. Essentially, it traversed through the word in the word state, outputting its input until it fell onto one of five symbols: /s/, /z/, /ɪ/, /t/, or /d/. In this case, the machine would give no output and go to state *s, z, ɪ, t,* or *d,* respectively. For the *s* and *z* states, if the following character was a word boundary #, we would return "# PRES", essentially stripping the present tense suffix from the verb, and replacing it with "PRES." A similar concept applied to the *t* and *d* states. If the following character was a word boundary #, we would return "# PST." The *ɪ* state was more intricate as this phoneme is used as a buffer between alveolars in the past tense, between sibilants in the present tense, and as part of the progressive suffix. For this state, we passed into the *ɪd* state if the following character was /d/, into the *ɪz* state if the following character was /z/, and into the *ɪŋ* state is the following character was /ŋ/, outputting nothing in all cases. From these three states, we searched to see if the following symbol was a word boundary # and output "# PST/PST.PTCP", "# PRES", or "# PROG", respectively.

If a word boundary was not detected in any of the eight aforementioned states, we would move to the appropriate states (*s* if the next character was /s/, *d* if the next character was /d/, *ɪ* if the next character was /ɪ/, *t* if the next character was /t/, *d* if the next character was /d/, or *word* otherwise), and output the name of the state it came from.

3.4 Orthographic-to-IPA Dictionary Mapping

We constructed two dictionaries. The first dictionary included the orthographic representation of select verbs of English as the entry. Each entry had five definitions: the bare form, the past, the progressive, the present, and the participle of the verb, all represented in IPA. This dictionary was used to test the performance of our forward model.

The second dictionary was purely defined in English orthography, following the same structure as the IPA dictionary. This dictionary contained the same 269 verbs as the IPA dictionary, with the English spellings instead of IPA transcriptions as definitions. The idea was to standardize IPA spellings and to map concrete English spellings to loose transcriptions in IPA. Due to dialectal variation and varying conventions, IPA transcriptions are somewhat unstandardized. For example, some dialects of English pronounce the word "aunt" as /ænt/, while others pronounce it as /ɑnt/. Additionally, some conventions use the /r/ while others use /ɹ/ to represent the English rhotic seen in words like "run." The implementation of this dictionary acts as a control in our testing such that each English verb maps to one IPA representation.

3.5 Dictionary Integration and Bidirectional Testing

These two dictionaries allowed the user to input English orthography for the words defined in our dictionary instead of requiring only IPA. Using the English dictionary, we were able to map user input to our standardized IPA spelling by locating its entry and which tense it is under, and finding the equivalent in the IPA dictionary. It is important to note that we do not store the tense information during the conversion to produce our outputs. Similarly, the incorporation of both dictionaries also allowed for seamless testing for both the inflector system and the inverse inflector.

3.6 Evaluation Methodology

To evaluate the performance of our Final State Transducers, we implemented a custom testing program designed to measure the accuracy of three systems: our individual rule-based FSTs, our Total FST (a words-and-rules hybrid model), and our morphological analyzer. Because FST behavior depends heavily on the phonological symbols used, we created a standardized IPA dictionary based on American English IPA transcription. This dictionary contains 269 verbs in total, consisting of 146 regular verbs and 123 irregular verbs. For each verb, the dictionary provides the bare form as well as the fully inflected forms for the present 3rd singular, progressive, past, and perfect tenses. During testing, we indexed this dictionary by verb and extracted the bare IPA form, which was tokenized and supplied as input to each FST. For the regular FST, we only evaluated the regular finite state transducers corresponding to our regular English morphological rules. In the case of the Total FST, the bare form fed into the transducer triggered an internal if/else decision structure: if the verb appeared on our hard-coded list of irregular verbs, the system applied the stored irregular mapping; if not, the input was routed through the regular verbal morphology FST. For every verb–tense pair, the predicted output of the transducer was compared directly against the corresponding gold-standard IPA form from our dictionary. Because FST outputs may contain internal markers or multiple possible paths, we used a normalization function that stripped off non-surface symbols (such as boundary markers) and collapsed the token stream into a single IPA string before comparison. The testing script incremented a running total of predictions and correct matches for each of the four tense categories, allowing us to compute accuracy as a simple proportion:

$$\text{Accuracy} = 100 \times \frac{\text{number of correct outputs}}{\text{total number of outputs}}$$

3.7 Morphological Analyzer Evaluation

The testing program for our morphological analyzer checks for two criteria of correctness. First, it extrapolates the base IPA form of the verb produced by the morphological analyzer and checks it against the bare form of the verb in the dictionary to ensure the outputs from our FSTs are correct. Then, it checks the morpho-syntactic tag produced by the analyzer against the expected substring corresponding to the tense being evaluated (for instance, the tag for a past-tense form must include PST) from the dictionary. Each tense is evaluated independently, and the corresponding counters track the total number of word forms and correct outputs. The correct count is only incremented when both the tag and the base words are correctly produced. After the verbs are processed, the accuracy is calculated by the above formula, giving us evaluation metrics for the tense labeling of verbs.

3.8 Evaluation Summary

These procedures ensure that every system is evaluated on an identical and phonetically consistent dataset, and that both regular and irregular verbs contribute to the final accuracy figures according to their frequency in the lexicon. The tester also logged each prediction, along with the raw FST output and the expected dictionary form, enabling qualitative inspection of errors. Together, this testing framework provided a transparent, reproducible method for assessing how well our FSTs captured the morpho-phonological processes of English verb inflection.

4 Results

The results of our evaluation provide a comparative overview of the performance of the Regular FST, Total FST, and the morphological analyzer across four key verb tenses. The Regular FST, which relies solely on morphological rules, shows high accuracy for present and progressive forms but struggles with past and perfect forms, reflecting the limitations of rule-based modeling for irregular verbs. In contrast, the Total FST, which integrates both regular and irregular transducers in a words-and-rules framework, achieves consistently high accuracy across all tenses. The morphological analyzer also demonstrates strong performance, effectively identifying the bare verb forms and correctly tagging inflections, highlighting the robustness of the system in parsing and analyzing verb morphology.

4.1 Tables of Results

Regular FST Model Accuracy Table

Tense	Accuracy Rate
Present 3rd Singular	91.8%
Progressive	94.8%
Past	56.9%
Perfect	56.9%

Table 1: Final accuracy report of the Regular FST, formulated on morphological rules.

Total FST Model Accuracy Table

Tense	Accuracy Rate
Present 3rd Singular	96.3%
Progressive	94.8%
Past	94.1%
Perfect	94.1%

Table 2: Final accuracy report of the Total FST, integrating both regular and irregular FSTs, mimicking a words-and-rules model.

Reverse Inflector Accuracy Table

Tense	Accuracy Rate
Present 3rd Singular	90.3%
Progressive	94.8%
Past	88.1%
Perfect	88.1%

Table 3: Final accuracy report of the reverse inflector, analyzing verb tagging and extraction of the bare verb.

5 Discussion

5.1 Progressive Tense Results

Our results corroborate our initial hypotheses. Based on Tables 1, 2, and 3, we observe that the progressive tense maintains consistently high accuracy across all models. This outcome is expected, as the progressive form lacks irregularities and mostly follows the morphological rules defined in our FST. Even when we implement a combined words-and-rules approach in the Total FST, the progressive output remains largely unchanged. The reason it does not reach 100% accuracy is likely due to modal verbs, which do not have progressive forms. In such cases, the FST may still attempt to generate outputs like kʊdɪŋ (coulding), following the progressive FST rules; even if irregularity conditions were specified, the output would be empty since there is no progressive equivalent for modal verbs.

5.2 Regular FST Results

For the FST implementing only a rule-based model, present tense verbs still show relatively high accuracy at 91.8%, since most follow standardized morphological rules with few irregularities. However, past and perfect tenses exhibit substantially lower accuracy, at 56.9% for both. This is because these tenses share the same underlying morphological structure, so any irregularity affecting the past also affects the perfect participle. Furthermore, these tenses contain a greater number of irregular verbs that are not captured by the rule-based FST, resulting in a high rate of errors. For example, the verb bring has the IPA transcription bɹɪŋ. When processed by the FST, the present tense rules correctly generate bɹɪŋz, while for the past and perfect tenses, the FST outputs bɹɪŋd, reflecting a regularized handling of these forms even though bring is actually irregular. This illustrates how the rule-based FST applies standardized morphological patterns, and highlights why some inaccuracies persist, particularly for irregular verbs in past and perfect tenses.

5.3 Total FST Results

In Table 2, where we integrate both regular and irregular FSTs, the Total FST demonstrates improved accuracy across the board. Present tense accuracy rises to 96.3%, reflecting the advantage of a words-and-rules approach, which allows the FST to distinguish between regular and irregular verbs. The most significant improvements are observed in the past and perfect tenses, where accuracies increase to 94.1%. This improvement arises from explicitly encoding irregular verbs into the FST, enabling it to handle a broader range of lexical items. In the case of bring, although the regular FST would output bɹɪŋd for the past and perfect forms, the total FST includes a specific arc for bring that correctly produces bɹɔt (brought), remedying the issue of irregularities. This demonstrates how incorporating irregular forms into the FST allows the model to handle exceptions, improving accuracy for verbs that do not follow standard morphological rules.

5.4 Reverse Inflector Results

In Table 3, the reverse inflector shows a slight decrease in accuracy, with present tense at 90.3% and past and perfect tenses at 88.1%. This drop is due to the analyzer's additional task of extracting the bare verb, which requires stripping inflectional endings and mapping the English verb form to its IPA representation. Errors can arise from ambiguities in the IPA transcription, incomplete handling of irregular verbs, or conflicts when multiple analyses are possible, leading to slightly lower accuracy compared to the Total FST. Despite this, the morphological analyzer still demonstrates strong performance, effectively capturing the majority of verb forms across tenses.

5.5 Comparative Performance Across Models

Taken together, the results demonstrate a clear hierarchy of system performance. The Regular FST performs well only in domains governed by highly productive morphological rules, while its accuracy deteriorates substantially in the presence of irregularity. The Total FST consistently achieves the highest accuracy across all tenses, supporting the effectiveness of a hybrid words-and-rules architecture. The reverse inflector performs slightly below the Total FST due to the inherent difficulty of inverse mapping, but still achieves robust accuracy across tense categories. These findings suggest that explicit storage of irregular forms, combined with rule-based generalization, is crucial for accurately modeling English verb morphology.

6 Limitations

6.1 Limitations of Accuracy Metrics

Here, our accuracy levels are quite high, but this does not necessarily mean that our FST perfectly represents the English language and its IPA transcriptions. The observed accuracy is highly dependent on the construction of our dictionary, and it is possible that the selection of words skews the results toward higher performance, particularly because most of the irregular verbs had already been hardcoded. This limitation must be kept in mind, and to further generalize and make our FST more robust, additional irregular verbs would need to be encoded. Previously, during the presentation, our accuracy levels were noticeably inflated, with 69% accuracy for past tense and 53% for participles. These figures were based on a much narrower dictionary of only 50 verbs, which contained a higher proportion of regular past-tense verbs, thereby inflating the past-tense accuracy. Additionally, the participle forms had been treated as identical to the past tense, even though they are distinct in some scenarios, a limitation that was later corrected while expanding the dictionary with more verbs.

Our model does not have a complete dictionary that allows the user to input any verb in English spelling and limited our ability to test the verbs of the English language. Although there is no universal consensus, it is estimated that the average English speaker knows 4,000 infinitive verbs, so our dictionary covers a very small proportion of those verbs. This also skews our results because the ratio of regular verbs to irregular verbs that we have included in the dictionary is unlikely representative of the actual ratio found in English.

6.2 Reverse Inflector and Sequence Parsing

During our implementation and testing phases, we discovered a few limitations of our model. For example, our reverse inflector continued to fail on the input #ʌntaɪ# (untie) and other verbs ending in /ɪ/ in their infinitive form. For example, the present tense input #ʌntaɪz# (unties) produces #ʌnta# as the bare form. This is because our current model cannot distinguish which parts of the sequence (ɪz# in this case) belong to the root and which parts belong to the suffix. It always assumes that the entire sequence is the suffix and removes it, which is why we see #ʌntaɪz# → #ʌnta#. Had we been given more time for development, we could have created a separate state to handle this sequence in a wider scope. Specifically, if we had detected a sibilant previous to this sequence, we could have been more confident that "ɪz#" was the suffix. Otherwise, we could have concluded that only "z#" was part of the suffix and appropriate to remove. We could have also applied this same concept to the past tense and alveolars /t/ and /d/ when dealing with ɪd# sequences.

6.3 Syllable Boundaries and Stress

Moreover, our current model ignores syllable boundaries and stress markers because of changes in these internal structures observed after affixation. Some words experience a shift in syllable boundaries across phonemes or in which syllable receives stress. For this reason, we decided to keep our IPA representations simple to ensure that there are no subtle inaccuracies.

6.4 Homophones and Ambiguities in IPA

Additionally, another limitation is the model's inability to distinguish between homophones in IPA. For example, the verb "ring" and "wring" are both represented by

#ɹɪŋ#. However, the past tense of "ring" is "rang" while the past tense of "wring" is "rung." Our model is only able to take one into account. In this specific case, it only accounts for "ring."

6.5 IPA Standardization and Source Errors

As previously mentioned, the lack of standardization of IPA representations served as a limitation because it required us to create our own standardization that the model works with. There was variation in representations among us and other sources that provided IPA translations, such as Wiktionary.org (Wiktionary).

Additionally, some of the sources we used to build our dictionary and irregular arcs contained errors. The source from Gallaudet University included several inaccuracies, where verbs that are orthographically irregular are not phonologically irregular, or provided the wrong participles (Gallaudet University).

7 Implications

As mentioned, our implementation follows Pinker's "words and rules" model, in which the parser searches for a verb in the irregular inventory before applying regular rules to produce the output (Pinker, 1998). This is exactly how our model functions: it first searches for the verb in the irregular inventory and, if it is not found, traverses through the regular FST to apply the morphological rules. The irregular inventory in our model also closely mirrors the way irregular forms are stored in the mental lexicon. Just as we needed to hard code irregular verbs for the model to produce correct outputs, humans must commit irregular forms to memory in order to produce them accurately.

Our work directly relates to Pinker's Words and Rules model, which suggests that regular forms are produced by rule application while irregular forms are stored and retrieved from memory (Pinker, 1998). Analogously, our Total FST first searches an irregular inventory for a verb and only applies the regular FST rules if no irregular entry exists. This design mirrors the cognitive distinction Pinker describes, providing a computational implementation of the words-and-rules hypothesis and allowing us to observe how such a system handles both predictable regular patterns and idiosyncratic irregularities.

Conclusion

Our implementation demonstrates that a hybrid words-and-rules approach effectively models English verb morphology, capturing both regular patterns and irregularities. The Total FST consistently outperforms the purely rule-based FST, particularly in past and perfect tenses, emphasizing the importance of storing irregular forms separately. The morphological analyzer performs robustly but exhibits slightly lower accuracy due to the additional task of extracting base forms. Limitations remain in handling homophones, incomplete verb coverage, and IPA standardization, suggesting future work should expand the verb lexicon and refine phonological representations. Overall, our study provides a computational framework for analyzing verb inflection in American English and supports cognitive models that distinguish between rule-governed and memorized linguistic knowledge.

References

Gallaudet University. Irregular verb list. https://gallaudet.edu/student-success/tutorial-center/english-center/grammar-and-vocabulary/how-to-use-verbs/irregular-verb-list/. Accessed: 2025-12-12.

International Phonetic Association. Full ipa chart. https://www.internationalphoneticassociation.org/content/full-ipa-chart. Accessed: 2025-12-12.

Gary F. Marcus, Steven Pinker, Michael Ullman, Michelle Hollander, T. John Rosen, and Fei Xu. 1992. *Overregularization in Language Acquisition*, volume 57 of *Monographs of the Society for Research in Child Development*. University of Chicago Press. PMID: 1518508.

Steven Pinker. 1998. Words and rules. *Lingua*, 106(1–4):219–242.

Simple English Wiktionary. Category: Verbs. https://simple.wiktionary.org/wiki/Category:Verbs. Accessed: 2025-12-12.