Predicting Future Phonological Changes Of Mandarin Chinese

Peijie Guo

1. Abstract

Linguists made historical reconstructions extensively based on existing patterns and common trends of phonological changes without the aid of sound data, yet previously little was done to extend such patterns into the future as a result of mostly human work in the field of historical linguistics. We created a system of translating IPA representation to vectors that captures characteristics of phonemes such as their articulatory location, and experimented with several machine learning models to capture existing trends from Old, Middle to Modern Chinese (Mandarin). Then we attempted to predict future Mandarin pronunciations based on existing phonological data and ancient reconstructions. The models were successful at capturing trends of Ancient to Modern Chinese shifts, as a great majority of the validation results were exactly correct. The predictions from the models vary in consistency of outputs and phonological inventory, but they are largely logical and may offer insight into future shifts. Therefore it is possible to predict phonological changes with machine learning, yet the accuracy of the predictions still needs validation in the future. If the predictions are proven to be effective in the future, the solutions may provide a method of capturing phonological changes and validating previous ancient reconstructions.

2. Introduction

While the future of languages may seem completely random, we can make conjectures about the phonological aspects of future languages. Linguists make reconstructions of ancient languages based on comparative methods and common patterns they notice in language shifts, despite the fact that there were no voice recordings. Given the same principle, it might be also possible to predict future phonological changes, given common patterns of shifts in the language and also comparison between several periods.

The main goal of this research is to attempt to predict the future phonological changes of Mandarin. However, an underlying objective is to validate the existing methods that linguists use to reconstruct ancient pronunciations. While incorrect predictions do not discredit the work done by historical linguists, correct predictions may prove the effectiveness of such a system and strengthen our understanding of ancient languages. This may also increase the role of computer technology in the studies of historical linguistics, which has traditionally been done by human linguists.

3. Background

Chinese was chosen for this task for its highly regular syllabic structure and consistency of one syllable per character ratio. On the other hand, despite its advantages, the downsides of ancient Chinese are that the written records are in logographic characters, and there is no assistance in orthography. Linguists have resolved these difficulties through poems and written phonological records such as *qieyun* and *guangyun*.

There are currently several systems of Ancient Chinese reconstruction. From the earliest complete reconstruction by Bernhard Karlgren to more modern reconstructions of Zhengzhang Shangfang and William H. Baxter, Ancient Chinese was extensively studied in the last century, and so is the field of historical linguistics (Karlgren 1).

On the other hand, little has been done on the subject of future linguistic shifts, perhaps as a result of the highly speculative nature of this topic. Yet the underlying basis has already been experimented with. We based the study on a system of transformation from phonetic representations in International Phonetic Alphabet (IPA) to vector representations and also from vectors back to phonetic representations. Prior to this research, there were already endeavors to tackle other problems in a similar manner. In particular, the Python module Panphon has shown an increased accuracy in NLP tasks when natural language information is recognized with phonetic vector representations, as opposed to character-based models (Mortensen et al. 3483).

However, this approach does not fit the nature of this problem, for it does not make distinctions between certain phonemes in that even with all the features, the module may still output multiple different phonemes, making mass transformation from vector to IPA difficult; and it also fails to identify certain combinations of phonemes, as they were not present in the training data for the module. In a study that attempts to predict shifts and expansions of pronunciations, a system with only a limited dataset may impede us from obtaining accurate results.

Apart from the necessity of establishing a one-to-one correspondence of each vector to each phoneme, the approach taken in the study also attempts to incorporate positional characteristics of phonemes, as positional patterns are often important in historical linguistics, While in PanPhon, positional features are not present in the vectors. Therefore, we chose not to use this approach in the research.

4. Dataset

The research is almost entirely based on the reconstruction made by Baxter and Sagart, as made public along with the book *Old Chinese: A New Reconstruction*. The dataset contains around 5000 entries of Chinese characters, For each character, the dataset provides a pinyin pronunciation for modern/new Chinese (NC), an ASCII friendly notation for Middle Chinese (MC), and also an IPA representation for Old Chinese reconstruction (OC). The dataset also provides information such as the definition of characters, but that information is not necessarily pertinent to the study.

Although tones are mentioned in the dataset, they will not be considered in this study. One reason is that OC was a toneless language and there is no method to establish a correspondence in OC (Sagart 11). Also, very little is known about the true value of tones in MC, which makes it difficult to create an accurate numerical representation of tones in MC (Wang and Sun 84).

The IPA system is the basis of this study, for it is recognized as the standard of phonological analysis, and also because of its ability to represent sounds beyond those within a set of languages. Thus, the pronunciations for Modern Chinese and Middle Chinese were first translated into IPA.

We used the python module Epitran to obtain the IPA pronunciations from pinyin. Middle Chinese IPA representations, on the other hand, require some translation work. However, this system, as described by Baxter and Sagart, is only intended to represent information given by sources, rather than serving as actual reconstructions (Baxter and Sagart 12). Therefore, the effectiveness of the translation from the notation to IPA is still uncertain. According to an earlier reconstruction of MC made by Baxter with a similar notation, there are 8 vowels in Middle Chinese inventory, four of which are represented with a different notation than IPA: + -> /ɨ/, ea -> /ɛ/, ae -> /æ/, o->/ʌ/ (Baxter 61). The initial and final consonants are replaced with IPA symbols indicated on the chart below based on the descriptions given in the books *Old Chinese: A New Reconstruction, A Handbook of Old Chinese Phonology* and *the Oxford Handbook of Middle Chinese Linguistics.*

Middle Chinese initial consonant notation to IPA (Baxter and Sagart 15-16; Wang and Sun 85)

Labials	p(h)->/p(ʰ)/	b->/b/	m->/m/			
Dentals	t(h)->/t(ʰ)/	d->/d/	n->/n/			
Lateral					l->/l/	
Retroflex stop	tr(h)->/ʈ(ʰ)/	dr->/ɖ/	nr->/ɳ/			
Dental sibilants	ts(h)->/ts(ʰ)/	dz->/dz/		s->/s/	z->/z/	
Retroflex sibilants	tsr(h)->/ʈʂ(ʰ)/	dzr->/ɖʐ/		sr->/ʂ/	zr->/ʐ/	
Palatals	tsy(h)->/tɕ(ʰ)/	dzy->/dʑ/	ny->/ɲ/	sy->/ɕ/	zy->/ʑ/	y->/j/
Velars	k(h)->/k(ʰ)/	g->/g/	ng->/ŋ/			
Laryngeals	ʔ->/ʔ/			x->/x/	h->/h/	

Middle Chinese coda consonant notation to IPA (Baxter and Sagart 61)

Zero->/∅/	w->/w/		j->/j/
ng->/ŋ/	wng->/ŋ/	m->/m/	n->/n/
k->/kʔ/	wk->/kʔ/	p->/p/	t->/tʔ/

Syllables that end in k, t, and p are known as *rusheng,* or entering tone. Although they are usually notated with diacritics such as /k̚/, we translated them as /kʔ/ for easier input. We also did not make distinctions between consonants and their respective combinations with the glide w, because those variants are not substantial in all periods of Chinese for it to increase another layer of complexity.

Having obtained the IPA pronunciations of three periods of pronunciations, we attempted to find obvious trends across all three periods. A few visualizations turned out to provide an interesting perspective.

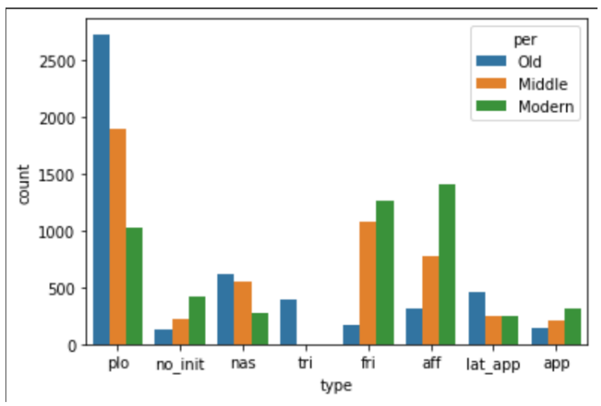


Figure 1

Figure 1 shows the distribution of manner of articulation across initial consonants of three periods. Initials with plosives show a steady decrease over time, and affricates, approximants, and null initials show a steady increase over time. There are some trends in nasals, approximants, fricatives, and lateral approximants, but they are not as steady. Trills ceased to exist in Middle Chinese.

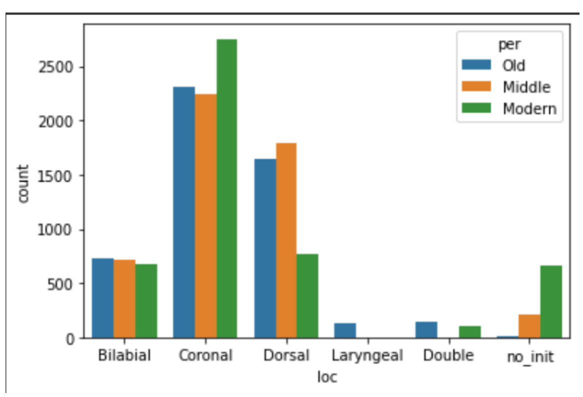


Figure 2

Figure 2 shows the distribution of articulation locations of the initial consonants. This may not be the best representation of the data as these categories are too broad to give specific details about the articulatory locations. Despite this, we can still notice that Bilabial initial consonants have maintained a stable presence. There are not as many trends with Coronal, Dorsal, Laryngeal and doubly articulated consonants, but there is a rise of syllables without an initial consonant.

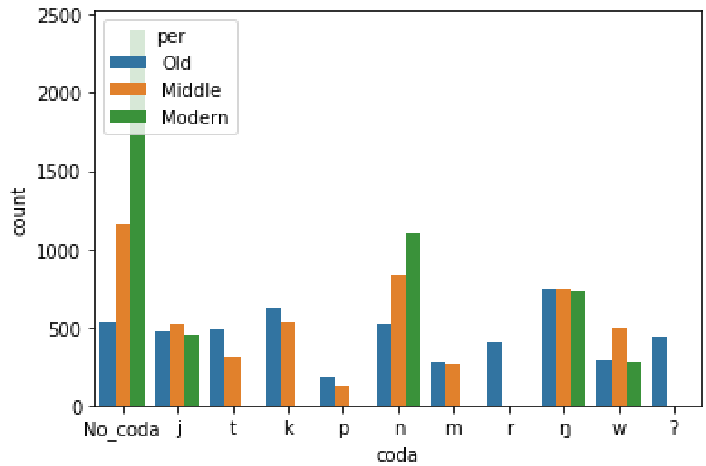


Figure 3

As there is a much greater restriction of available phonemes in coda final consonants of a syllable, Figure 3 shows the distribution of individual coda consonants. the velar nasal /ŋ/ is very stable in its presence. The palatal /j/ final also has few variations. The alveolar nasal /n/ has a steady increase over time. In Mandarin, syllables that end in a stop, indicated by /t/ /k/ /p/ /ʔ/, and also nasal /m/, no longer exist, but they are largely conserved in many other dialects such as Cantonese. The missing stops explains the spike increase of syllables with no coda consonants. The trill /r/ doesn't exist since old chinese and the approximant /w/, which may also be notated as a non syllabic /u̯/, does not exhibit a clear pattern.

Because vowels are harder to visualize as a whole without venturing into higher dimensions, we broke down the characteristics of the vowels. We focused exclusively on monophthongs as trends in diphthongs are difficult to visualize and more difficult to observe. Figure 4 shows the frontness of monophthongs, while Figure 5 shows the height of monophthongs.

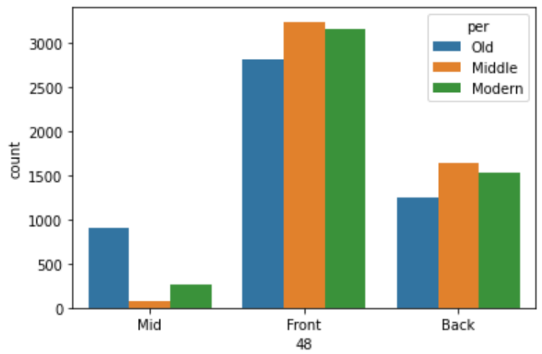


Figure 4

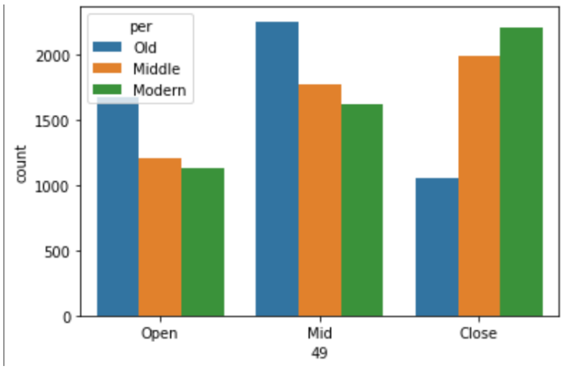


Figure 5

Vector transformation

Chinese has a highly regular syllabic structure. According to reconstructions of many linguists, the maximal structure of an old chinese syllable contains initial consonant, medial (or glide), vowel and a coda, with optionally added affixes, also called pre-initial or post-coda (Sagart 14). The combination of the vowel and the coda is also referred to as the rime, and when combined with the medial, they make up the final of a syllable. Middle Chinese retained the basic structure of the four principle parts but lost the affixes, and such features were then passed on to modern Chinese. Regardless of the dialect, a modern Chinese syllable also contains these four parts (Duanmu 2). Such consistency in phonological features allowed for more potential approaches to the problem while simplifying the process of transformation. While there are still heated debates about whether Chinese is an isolating language (Packard 355), it is evident that each character only corresponds to one syllable, as opposed to other languages, where insertion and deletion of a syllable from a word frequently happen. Thus a vector must be able to represent all possible sounds for all characters available in a dataset, along with potential future sounds, all within a certain length.

We produced an expert system to vectorize IPA representations for several reasons: The dataset is not sufficient for a computer developed vectorization algorithm, the vector must also extend into representations not seen in the datasets, and the models must be able to detect specific patterns in shifts of articulatory locations, vowel placements, aspirations, etc., and that may not be as effective with a system such as one hot encoding or other machine learning based vectorization algorithm.

The resulting vector was one of dimension 86, and is composed of the six principal parts of a Chinese syllable—the pre-initial, initial, medial, vowel, coda, and post-coda. Consonant vectors are each of dimension 16, example shown below:

/m/ -> [1, 1, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1]

The first two features of the vector indicate the articulatory location, as organized by the IPA standards, of the phoneme, ranging from 1 to 12 (1-bilabial, 2-labial dental, 3-dental, 4-alveolar, 5-post alveolar, 6-retroflex, 7-palatal, 8-velar, 9-uvular, 10-Pharyngeal, 11-Glottal, 12-null). Values below 1 or above 11 are considered to be null /∅/. The first two features of pulmonic consonants are always the same, while affricates and voiced labial velar approximants /w/ receive different markers in the first two features to indicate different articulatory locations. In this manner, the vector shows shifts in articulation.

The remaining features of the vector are organized as categorical data. The next 8 features marks the manner of articulation (nasal, trill, tap or flap, fricative, lateral fricative, approximant, lateral approximant,affricate), and the remaining 5 features indicates other characteristics of the consonants: aspiration, pharyngealization (only present in Old Chinese), glide, stop, and voice.

As an example, the phoneme [ts^h] is represented by (4, 4, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 0, 0, 0, 0). The initial 4’s indicate that the consonant is alveolar, 1 at the 11th place indicates that the consonant is an affricate, and 1 at the 12th place indicates that the consonant is aspirated.

Vowels are represented by 3 components. The first component describes the backness and ranges from 0 to 20. The second component describes the height of the vowel and ranges from 0 to 30, as shown on Figure 6 (International Phonetics Association). The third component describes the roundness, for which 0 indicates unrounded and 5 rounded. For processes where a standard scale does matter, we normalized all the components before performing those processes. The following chart indicates the values of backness and height. Therefore, the vowel /y/ is represented by the vector [0,30,5], and the vowel /ɐ/ is represented by the vector [10,5,0]. Although the vowels /ɐ/ and /ə/ are not defined for roundness in IPA, they are pronounced in various Chinese dialects as unrounded, such as /ə/ in Mandarin and /ɐ/ in Cantonese (Lee and Zee 110). For such reasons the roundness value for these vowels are designated to their respective preferred pronunciation in Sinitic languages.

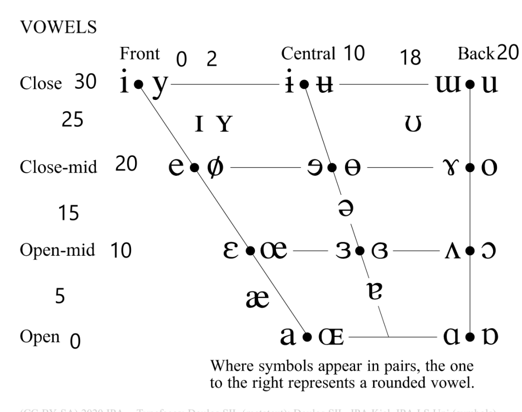


Figure 6

The maximal vowels present in one chinese syllable are restricted to dipthongs. In certain notations, semivowels are notated as /i̯/ or /u̯/ and certain characters are written with a triphthong notation. For the consistency of vector representation, these semivowels are replaced with the respective approximants /j/ and /w/ representation. These transformed representations are then classified as the medial or the final, e.g. /t͡su̯ɔ/ -> /t͡swɔ/ or /ai̯/ -> /aj/. Triphthongs are thus effectively reduced to monophthongs.

Syllabic consonants have thus far only occurred in Mandarin Chinese, and there has not yet been a perfect way to represent them through a vector. The current solution is to replace a syllabic consonant with a vowel sound most similar to that of the syllable. There are many ways of notating syllabic consonants, the most common way the three syllabic consonants in Mandarin chinese are /z̩/ (and optionally /ʐ̩/) /ɻ̩/, and the syllabic parts are replaced by /ɨ/ and /ɘ/ respectively, e.g. /z̩/ -> /zɨ/ and /sz̩/ -> /sɨ/.

The six segments are then concatenated into one complete vector, as shown below:

觀 - Old Chinese pronunciation /c.qʷˤar-s/
Input syllable: /c.qʷˤar-s/
Initial final separation (c.qʷˤ, a, r-s)
Phoneme tokenization (c, qˤ, w, a, r, s)
Vector transformation ([7, 7, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0]
[9, 9, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0]
[1, 8, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 1, 0, 1]
[0, 0, 0, 0, 0, 0]
[4, 4, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1]
[4, 4, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0])
Vector Concatenation [7, 7, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 9, 9 … 0, 0, 0, 0, 0, 0]

The dataset provided makes distinctions of tightly attached pre-initials, such as /m/, and their loosely attached pre-initial counterparts, /mə/ (Baxter and Sagart 53). However, there were no such distinctions in previous reconstructions made by Baxter (Baxter 175-177), and affixes no longer exist since Middle Chinese. Thus, the vector makes no distinctions between the tightness of pre-initials to prevent overcomplication of features. For the same reason, when two pre-initials are attached to an Old Chinese, only the first pre-initial is marked in the vector, and the pre-initial/post-coda segment is only used to distinguish between variations of Old Chinese pronunciations.

The vectors are transformed back into IPA representations in a similar fashion. The complete vector representation is split into 6 parts in a similar format as the transformation from IPA to vector. To address regression outputs that are usually not integers, the consonant vectors are placed in a process of fuzzy phoneme matching: If either one of the first two features is under 1 or over 12, the phoneme is deemed null. Then a difference is found between the two articulations values. When a threshold is exceeded, the pool of available phonemes is restricted to the five sounds that have existed in Chinese: /tʃ/, /tɕ/, /ɕ/, /dʑ/, /w/. On the other hand, if the threshold is not exceeded, features that indicate manner of articulation are converted to one hot encoding, which is then used to limit the pool of available phonemes to only contain those that fit their manner of articulation. From the pool, one final consonant is chosen from calculating the euclidean distance between the available IPA symbols and the input vector. Modifiers such as aspiration marks are added on to the final consonant in a one hot encoding format. Vowel vectors are transformed through comparing the distance between all vowels and picking the closest IPA representation. Separate phonemes are then concatenated together, from which one complete IPA representation is obtained.

For example, when a consonant vector [1.1037 1.3289 1.0432 -0.0160 -0.0007 0.0023 -0.0703 0.0018 -0.0016 -0.0173 0.0090 0.8444 -0.0021 -0.0011 -0.0010 -0.0404] is inputted into the transformer, the transformer first verifies the first 2 indices were not close to 0 or over 11.5 if so, the output would be null. However, in this case, the two features were not 0 nor 11.5. Thus the algorithm evaluates the next 9 features and converts those features into categorical format. With this vector, the first feature was the greatest among the 11, which means that the phoneme is most likely to be a plosive and the available phonemes are restricted to only the plosives /p/ /b/ /t/ /d/ /ʈ/ /ɖ/ /c/ /ɟ/ /k/ /g/ /q/ /ɢ/ /ʔ/. Among these consonants, distance is calculated between the input phoneme and the available phonemes. The closest by distance is /p/, whose vector representation is [1, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0]. The last 5 features correspond to aspiration, pharyngealization, glide, stop, and voice. The feature of aspiration is greater than 0.5, thus the phoneme includes the modifier /^{h}/, and the algorithm outputs /p^h/.

With this system, the complete dataset now includes the characters and their respective pronunciations across all three periods, and also the vector representations for the three pronunciations.

5. Methodology/Models

Because of the speculative nature of this topic, there is no definitive method of training or validating the system. Therefore, we took several approaches to analyze the phonological changes across three periods.

Errors between two vectors are calculated through a custom mean squared error function. The vectors are first transformed into a closest IPA representation and then transformed back to the vector of the closest IPA representation to prevent non-zero errors on identical IPA pronunciations, because the models often output inexact vectors, when transformed, render to be correct. Both vectors are normalized, from which a mean squared error is derived.

Statistical model - OC + MC -> NC/ MC + NC -> FC

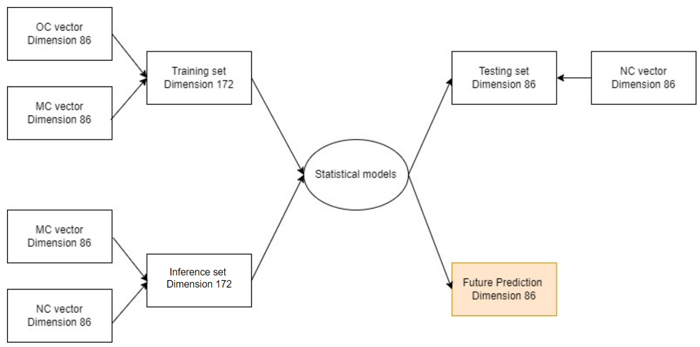


Figure 7

One way of validating such a system is to train and test the models on existing data. We first split the dataset into training, validation, and testing sets, from which input and output values are derived. The training set contains about 94% of the total 4967 character dataset, and the validation and testing set each contain about 3% of the total dataset. We concatenated OC and MC vectors to create an input vector of dimension 172 and left NC vectors as the output vector of dimension 86. We trained the models with the OC and MC vectors and tested them against the NC vectors, and then obtained future pronunciations with concatenated MC and NC vectors, as demonstrated on Figure 7.

We first experimented with several statistical machine learning models—decision tree, random forest, and extra trees regressor offered by Scikit learn. With the assistance of the Grid search algorithm, we determined that across all three algorithms, there is not one definitive optimal hyperparameter for each model. Based on the frequency of lowest loss scores, the max_depth of the decision tree regressor is set to 100, n_estimator and max_depth of both random forest and extra trees regressor are set to 200.

Neural network - OC + MC -> NC/ MC + NC -> FC

We then experimented with a multilayer perceptron model (MLP) and also a simple Long short-term memory (LSTM) model. The LSTM model was set up with 2 LSTM layers of 200 neurons and an output dense layer with 86 outputs. The MLP model was set up with 2 hidden layers, each with 300 perceptrons and an output dense layer with 86 perceptrons.

We normalized the vectors down to a scale of 0 to 1 for every component to fit the nature of the models. With that dataset, we reshaped the dataset to 2 steps of 86 components per character, and trained using OC and MC vectors as inputs with NC vectors as the next step. When predicting future pronunciations, we reshaped MC and NC vectors into 2 steps and fed the dataset into the model. We used the same technique of concatenating normalized OC and MC vectors as the training sets for MLP and feeding normalized MC and NC vectors to predict future pronunciations.

Statistical model MC -> NC/ NC -> FC

We also attempted to feed only MC vectors into the statistical training algorithm, with the same parameters as the previous statistical models. We used only the NC vectors as the input to get predictions, and the model returns future Chinese predictions based on the patterns of phonological changes from only MC to NC, as demonstrated on figure 8.

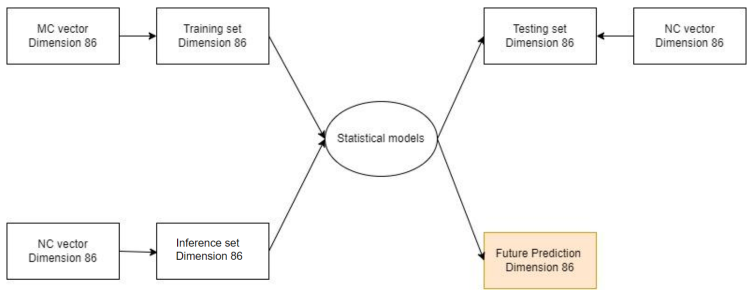


Figure 8

Polynomial model - OC + MC + NC -> FC

The last approach we attempted was a polynomial model. It is an experimental approach and is not as sophisticated as the other approaches. However, it may still be interesting to examine the results of the solution nonetheless.

This polynomial model aims to address the concern that the predictions may just be replicates of the previous phonological changes, rather than actual predictions. Therefore, this approach takes in vectors from all three periods and analyzes the vectors on a character basis and not the dataset as a whole. The model finds a curve of the least degree that captures the shifts from Old to modern chinese along a time axis OC corresponds to the time axis as 1, MC – 2, and NC – 3, and this is done through increasing the degree of coefficients until the mean squared error of predictions for all three time periods all reach 0. Because there is not a definitive period for each stage of development, training sets and also future predictions are not definitive on the time axis, but rather simple integers that indicate the sequence of pronunciations. We attempted to find predictions with several time values greater than 3.

6. Results and Discussion

Statistical model - OC + MC -> NC/ MC + NC -> FC

Across all three models, the extra trees regressor performed the best with a mean squared error of around 0.010. On the other hand, the decision trees regressor and random forest regressor had an average mean squared error of around 0.011. These three models were successful at validating almost all of the testing sets. A few samples of erroneous modern Chinese validations are shown below,

Char	OC	MC	NC	Tree	Forest	ExtraTrees
番	/pˤar/	pa	/pwɔ/	/pʰdʑɵɴ/	/pʰwɔ/	/pʰwɔ/
劾	/gˤək/	hok	/xɤ/	/xɤ/	/çɤ/	/xɤ/
紐	/nruʔ/	nrjuw	/njoʊ/	/njoʊ/	/nɰoʊ/	/nɰoʊ/
螘	/m-qʰəjʔ/	ngj+j	/i/	/ɕi/	/tɕi/	/i/
曬	/sre-s/	srea	/sa/	/sa/	/sæ/	/sa/

Most of the errors seemed to have come from initial consonants, whereas the vowels were quite accurate.

We then fed Middle Chinese and Modern Chinese pronunciation vectors into the models so that the models would predict a supposedly future pronunciation based on the previous shifts that it had observed. A small sample of predictions is shown below, all of which were from the testing dataset whose modern Chinese pronunciations predicted by the model have been validated to be correct.

Char	OC	MC	NC	Tree	Forest	ExtraTrees
累	/roj/	ljwe	/lej/	/li/	/li/	/ɭɪɰ/
緧	/tsʰu/	tshjuw	/t͡ɕʰjoʊ/	/tsʰwej/	/ʈʂʰdʑɪɰ/	/ʈʂʰʏ/
變	/pron-s/	pjen	/pjɛn/	/pɰæn/	/pæn/	/pɛan/
管	/kʷˤanʔ/	kwan	/kuan/	/ɕyan/	/ʈʂøan/	/tɕeæn/
祈	/c.ɢər/	gj+j	/t͡ɕʰi/	/ʈʂɨ/	/ʈʂʰɨ/	/ʈʂɨ/

Based on the testing samples, It is worth noting that despite having a greater error, decision trees regressor outputs results that were the most logical and probable, because the manner in which the algorithm is trained makes decision tree most probable that it adheres to the standard mandarin phonology. Although the outputs are more logical and easier to pronounce for Mandarin speakers, this type of algorithm is culpable of restricting the possibilities of expansion of pronunciations in the future.

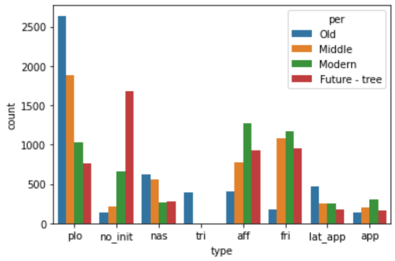


Figure 9

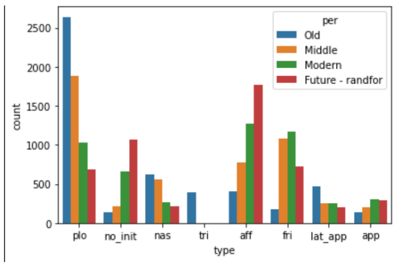


Figure 10

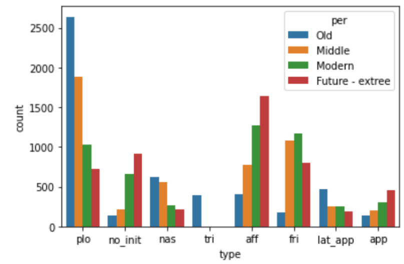


Figure 11

When compared with the trends in the dataset, all three models largely continued the trends in the existing dataset, and that shows, in theory, the effectiveness of the models. Trends in plosives, lateral approximants are quite consistent. Some trends were shown in nasal consonants and approximants, Syllables with no initials showed consistent trends with random forest regressor and extra trees regressor, whereas with decision tree, it experienced an interesting leap in numbers. The number of fricatives, however, experienced a decline across all three models despite increasing patterns, and affricates also experienced a decline in decision tree but the pattern continued in random forest and extra trees regressor. The remaining graphs for all models are included in appendix A.

Neural network - OC + MC -> NC/ MC + NC -> FC

Compared to the previous approach, the Neural network approach was somewhat less consistent with both the mean squared error and also the phonetic output. The models were trained with the Early Stopping method so that they don't overfit. The LSTM model trained for an average of 30 epochs. The following graph shows one instance of a training with 27 epochs.

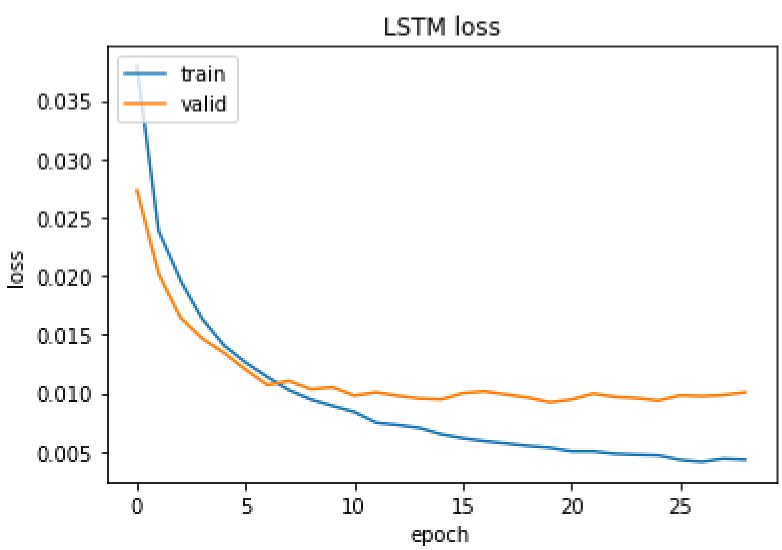


Figure 12

The model achieved a mean squared loss of around 0.010 with the testing dataset. The training and validating curve appeared to be reasonable and gives much confidence to the performance of the models. However, the model showed a sign of overfitting at the end of the training. On the other hand, the MLP models were trained for an average of 25 epochs with an early stopping method. The graph shows an instance of training with 28 epochs.

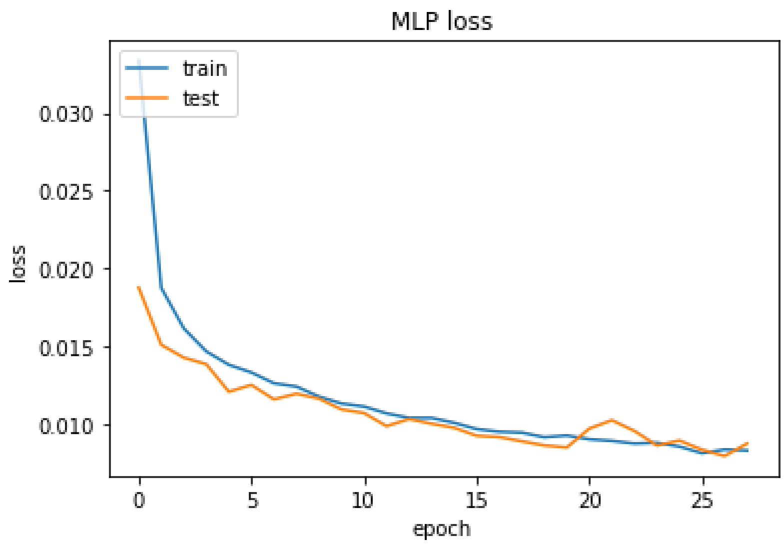
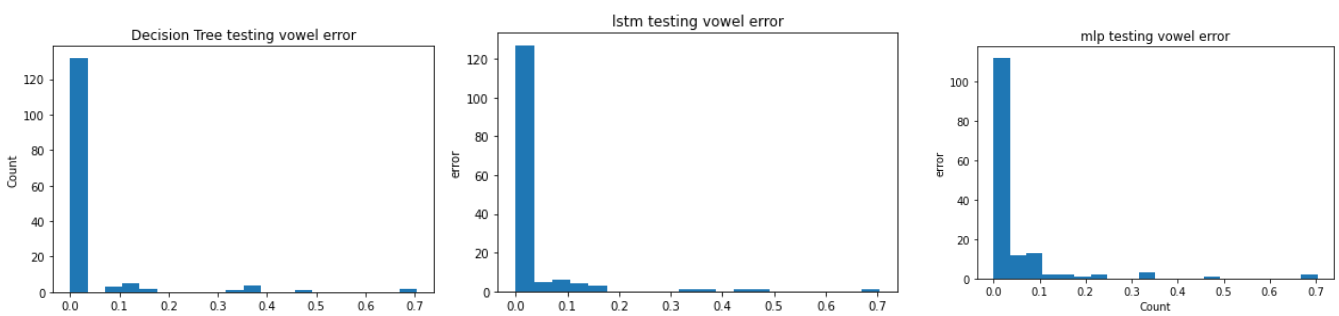


Figure 13

While the MLP model appeared to be much more consistent than LSTM on the training loss graph, the mean squared loss was significantly greater than any of the previous models at 0.016, but the error alone should not discredit the outputs of the MLP model, because future pronunciations should not simply replicate the shifts that already happened. Variations of modern Chinese predictions can still demonstrate other potential phonological changes in the future.



It is also interesting to note that based on the testing sample, the vowel errors produced by the neural networks are less significant than the testing data in that many erroneous phonemes differ by only a small feature, but they are greater in frequency. The chart below demonstrates a few samples of incorrect modern Chinese validations, which are far more frequent than the statistical models.

Char	OC	MC	NC	LSTM	MLP
繫	/cə.kˤe-s/	kej	/ɕi/	/ɕi/	/çi/
幅	/prək/	pik	/fu/	/fu/	/fuʊ/
搌	/trenʔ/	trjen	/ʈʂan/	/ʈʂæn/	/ʈʂjɛn/
屈	/nə-kʰut/	khjut	/tɕʰy/	/tɕʰʉ/	/ʈʂʰʉʊ/
揭	/kʰrat/	khjet	/tɕjɛ/	/tɕɪ/	/tɕʰjɪe/

A sample of future Chinese predictions is listed below. The NC pronunciations of the characters in the sample have either been predicted correctly or almost predicted.

Char	OC	MC	NC	LSTM	MLP
幣	/bet-s/	bjiej	/pi/	/pʰej/	/tsɪə/
浪	/rˤaŋ/	lang	/laŋ/	/ʈaŋ/	/laŋ/
啖	/m-rˤamʔ/	dam	/tan/	/ʈan/	/tan/
顛	/tˤin/	ten	/tjɛn/	/ʈʂjan/	/tsjɛn/
宜	/ŋraj/	ngje	/i/	/i/	/e/

Statistical model- MC -> NC/ NC -> FC

The outputs with only one set of input vectors are very similar to that with two vectors. The models performed with almost equal accuracy, with the mean squared losses for all three models being at approximately 0.010. A sample of erroneous predictions are shown below.

Char	MC	NC	Tree	Forest	ExtraTrees
縶	trip	/ʈʂɨ/	/tɕi/	/ʈʂɨ/	/ʈʂɤ/
劾	hok	/xɤ/	/xɤ/	/çəɰ/	/xɤ/
紐	nrjuw	/njoʊ/	/njoʊ/	/nɰoʊ/	/nɰoʊ/
碭	dang	/taŋ/	/tʰaŋ/	/tʰaŋ/	/tʰaŋ/
曬	srea	/sa/	/ʈʂaj/	/ʃæ/	/sa/

Compared to the previous statistical approach, where we only fed Middle Chinese vectors into the training, there is a slight increase in consonantal prediction errors. On the other hand, the vowel error rates are similar to the previous approach, as exemplified in the graphs below.

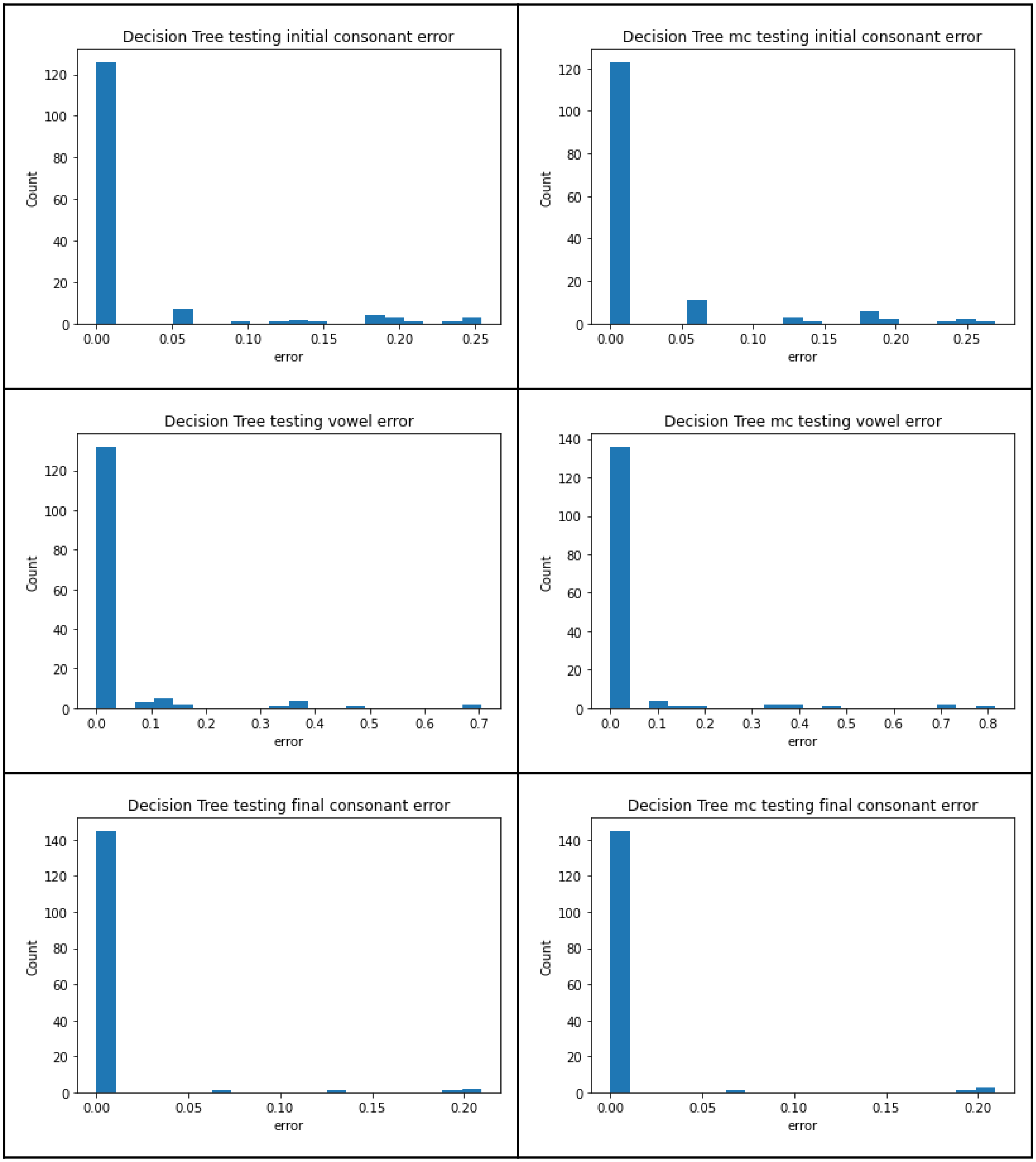


Figure 14

It also seems to be that models trained on MC vectors predict medials less consistently than models trained on both OC and MC vectors, as affricates show up in as medial consonants, which does not conform to Chinese phonology. Overall, these results were almost as consistent as the other approaches.

A sample of future chinese predictions is listed below

Char	MC	NC	Tree	Forest	ExtraTrees
累	ljwe	/lej/	/li/	/li/	/li/
緧	tshjuw	/t͡ɕʰjoʊ/	/ʈʂʰwej/	/ʈʂʰwej/	/ʂʊ/
變	pjen	/pjɛn/	/pɰæn/	/pan/	/pæn/
管	kwan	/kuan/	/tɕyan/	/tɕøan/	/ceæn/
祈	gj+j	/tɕʰi/	/ʈʂʰɨ/	/ʈʂʰɨ/	/ʈʂʰɨ/

Polynomial model = OC + MC + NC -> FC

There is a much greater uncertainty to the predictions made by the polynomial models. We could not evaluate the accuracy of this method because all curves are fitted specifically to each character. Some phonemes that were no longer used in modern Chinese show up in predictions as a result of the curve trying to fit a recurrent trend. Also, unlike the previous models, each component of the vector are fitted individually, which makes each phonemes unrelated to each other, which means that the phonemes and their features lack context in the syllable, and most importantly, the model predicts affixes based on the curves, even though that feature had long disappeared since Old Chinese. Therefore, affixes will not be considered into the predicted pronunciation when we compare the results from all the data.

A sample of predictions by the polynomial models is shown below. The 3 sets of predictions corresponds to the time values (3.2, 3.5, 4) inputted to obtain the results

Char	OC 1	MC 2	NC 3	Poly 3.2	Poly 3.5	Poly 4
理	/mə.rəʔ/	li	/li/	/lɪ/	/ɳ-rəʔ/	/rɑʔ/
幣	/bet-s/	bjiej	/pi/	/pɪik-x/	/peip-f/	/pai/
顛	/tˤin/	ten	/tjɛn/	/tɹɛn/	/tˤʋæn/	/tˤan/
碭	/lˤaŋ-s/	dang	/taŋ/	/taŋ-χ/	/lˤaŋ-s/	/lˤaŋ/

宜	/ŋraj/	ngje	/i/	/i/	/iɰ/	/eɹ/

Despite the unconventional phonology, the outputs can still give valuable information as they offer an insight into the general direction of the shifts, and also this method would, in theory, prevent a reduplication of phonological changes from MC to NC. Aside from affixes, pharyngealizations would also not be considered when we review the outputs for the same reason as affixes.

7. Conclusion

All the models predicted modern pronunciations quite accurately. However, because there is no definitive correct result for a prediction, and future phonological changes can occur in almost all directions, we cannot settle on one model that had the lowest loss scores, but rather, we must consider all the possibilities that the models give as if all of the predictions are equally as likely to happen. Even if one pronunciation is chosen, the results would only indicate one such shift that may occur in Mandarin, as Mandarin may eventually split into separate dialects.

If some of the predicted pronunciations are observed in the future, and they may come in the form of different dialects, predicting future phonological changes with machine learning can be proven effective and be used in other applications or other languages. Not only that, the greater goal of this research is to validate the accuracy of current reconstructions of Old Chinese and Middle Chinese, and also various methodologies that we utilize to obtain those reconstructions.

However, this research is nowhere complete, and it only serves as the first step in this field for other research ideas in the future. There are still many improvements in the experimental setup to be made. The scaling of components such as the vowel components is not consistent, leading to uncertainties regarding normalization of data during training. The threshold that distinguishes a doubly articulated consonant is also not optimized, as the number was derived from the smallest difference of a doubly articulated consonant. We did not have a good method to represent syllabic consonants. There is room for improvements with the neural models, as we did not place too much emphasis on them despite the fact that they have much more potential. The polynomial model still needs a great amount of optimizing for it to predict results as well as the other models do, while preventing duplication.

Perhaps the greatest challenge to fully realize this research question is that there are too many historical factors that can influence language change. Contact with another group is inevitable in the real world and with that comes unexpected shifts in phonology and syntax in a language, and currently there is no best way to predict such events in the future. With current solutions, we can model language shifts with existing phenomena (Kandler and Steele). However, this approach may still fall short if we have no knowledge of unexpected events in the future. Furthermore, the models predicted the results based on a character basis, whereas the pronunciations may also be influenced by other characters that make up morphemes. Therefore, we may never predict future translations perfectly.

We may consider how this research can extend into other languages. The main drawback of the current solution is that it restricts the syllabic structure. Unlike Sinitic languages, most other languages in the world are not restricted to one syllable per morpheme, but multiple syllables per multiple morphemes. The current system focuses primarily on investigating the shifts of single phonemes, and ignores all other phonological processes such as insertion and deletion of syllables, or shifts in tones. Another potential solution to extend such a method into other languages is to analyze individual syllables, as syllables tend to have a more consistent structure than words in synthetic languages. Despite these underlying problems, the models performed decently, and they output logical results for both validation and prediction, which prove that this approach can certainly be applied to further research and experiments.

8. Acknowledgements

I would like to express my gratitude to William Baxter and Laurent Sagart, whose reconstruction of Old Chinese and Middle Chinese served as the basis of this research. I also wish to show my appreciation to Kush Khosla and the InspiritAI organization for assisting me to make this endeavor possible.

9. References

Baxter, William H. A Handbook of Old Chinese Phonology. Mouton De Gruyter, 1992.

Baxter, William H. “The Baxter-Sagart Reconstruction of Old Chinese.” Ocbaxtersagart.lsait.lsa.umich.edu, https://ocbaxtersagart.lsait.lsa.umich.edu/.

Baxter, William H, and Laurent Sagart. Old Chinese a New Reconstruction. Oxford University Press, 2016.

Chollet, François, et al. ‘Keras’. N.p., 2015. Web.

Duanmu, San. “Syllable Structure in Chinese.” Syllable Structure, 2008, pp. 72–85., https://doi.org/10.1093/acprof:oso/9780199267590.003.0004.

Hunter, John D. “Matplotlib: A 2D Graphics Environment.” Computing in Science & Engineering, vol. 9, no. 3, 2007, pp. 90–95., https://doi.org/10.1109/mcse.2007.55.

International Phonetic Association. “Ipa Kiel / Ls Uni (SERIF) Font:” IPA English Charts, https://www.internationalphoneticassociation.org/IPAcharts/IPA_chart_orig/IPA_charts_E.html.

Kandler, Anne, and James Steele. “Modeling Language Shift.” Proceedings of the National Academy of Sciences, vol. 114, no. 19, 2017, pp. 4851–4853., https://doi.org/10.1073/pnas.1703509114.

Karlgren, Bernhard. “The Reconstruction of Ancient Chinese.” T'oung Pao, vol. 21, no. 1, 1922, p. 1., https://doi.org/10.1163/156853222x00015.

Lee, Wai-Sum, and Eric Zee. “Standard Chinese (Beijing).” Journal of the International Phonetic Association, vol. 33, no. 1, 2003, pp. 109–112., https://doi.org/10.1017/s0025100303001208.

Mortensen, David R. et al. ‘Epitran: Precision G2P for Many Languages’. Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018). Miyazaki, Japan: European Language Resources Association (ELRA), 2018.

Mortensen, David R. et al. ‘PanPhon: A Resource for Mapping IPA Segments to Articulatory Feature Vectors’. Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers. Osaka, Japan: The COLING 2016 Organizing Committee, 2016. 3475–3484.

Pedregosa, Fabian et al. “Scikit-learn: Machine Learning in Python” Journal of Machine Learning Research 12.85 (2011): 2825–2830. Web.

Packard, J.L. “Chinese as an Isolating Language.” Encyclopedia of Language & Linguistics, 2006, pp. 355–358., https://doi.org/10.1016/b0-08-044854-2/00149-8.

Sagart, Laurent. ‘The origin of Chinese tones’. Proceedings of the Symposium/Cross-Linguistic Studies of Tonal Phenomena/Tonogenesis, Typology and Related Topics. Tokyo, Japan: Institute for the Study of Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies., 1999. 91–104.

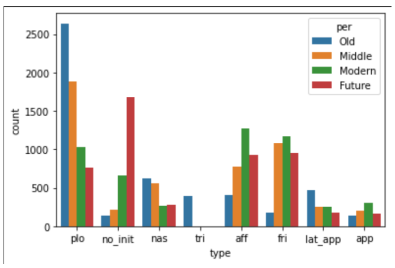
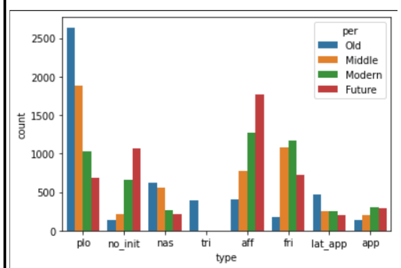
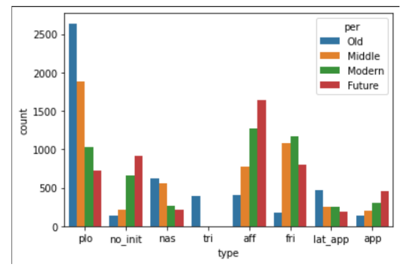
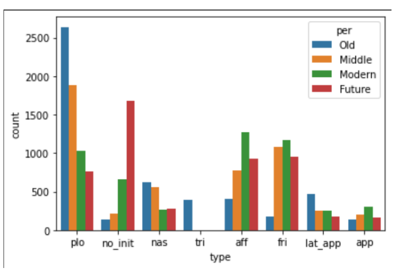
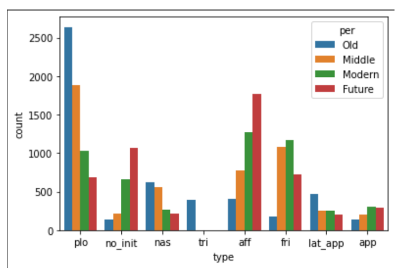
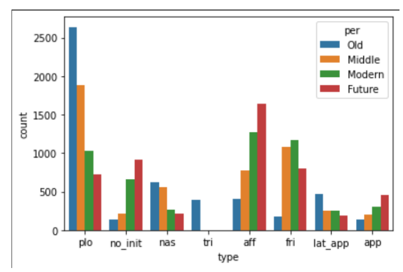
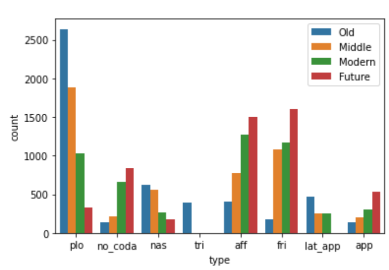
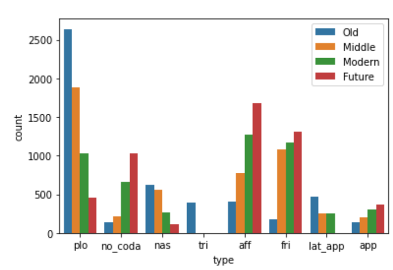
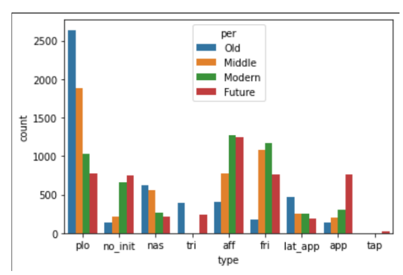
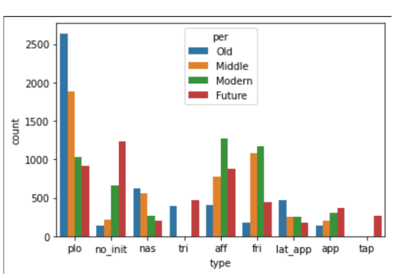
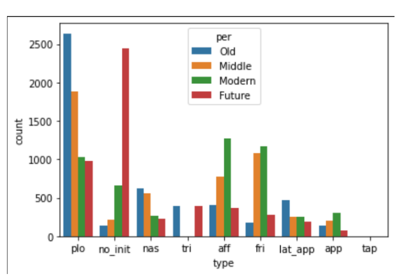
Sagart, Laurent. The Roots of Old Chinese, John Benjamins Pub. Co., Amsterdam, 1999, p. 14.

Wang, William S-Y, and Sun. The Oxford Handbook of Chinese Linguistics, Oxford University Press, New York, NY, 2015, p. 84.

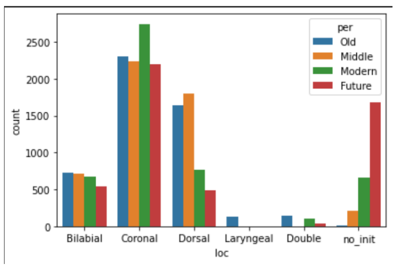
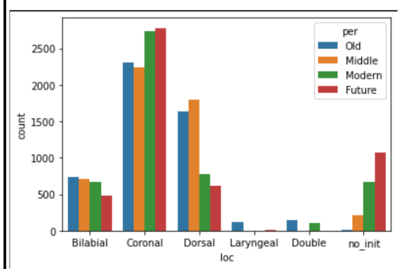
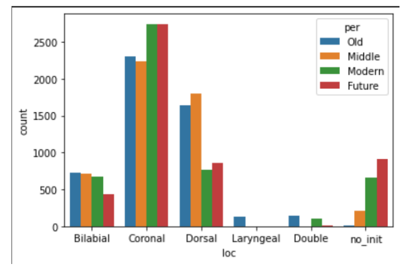
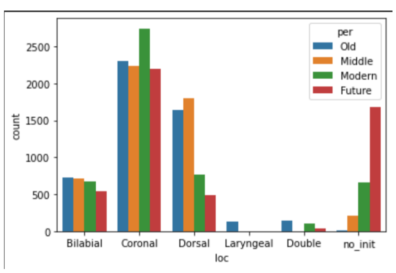
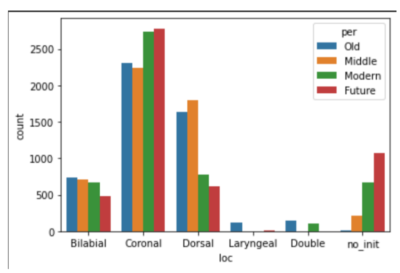
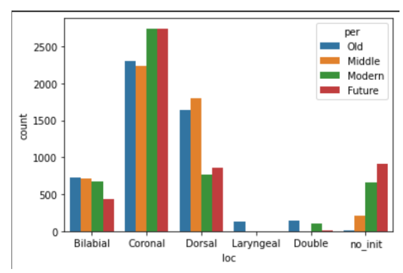
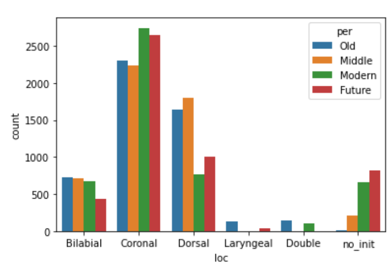
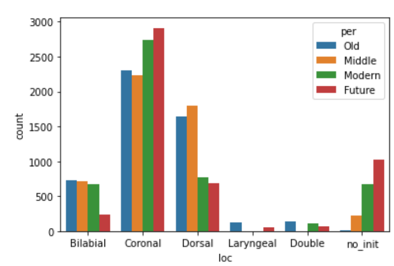
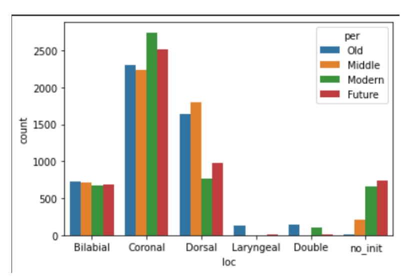
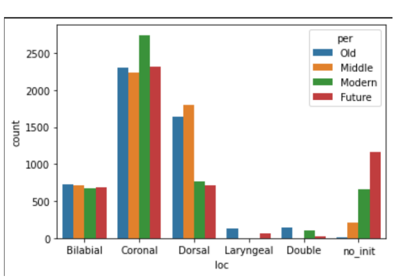
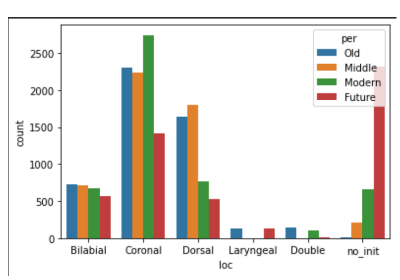
Zhong, Li. “Zui Chang Yong De Yi Qian Ge Han Zi Shi Yong Pin Lü Pai Ming” [Frequency Rank of One Thousand Most Used Chinese Characters]. Hua Yu Wang [Website of the Chinese Language], https://www.thn21.com/base/zi/17300.html.

10. Appendix A graphs

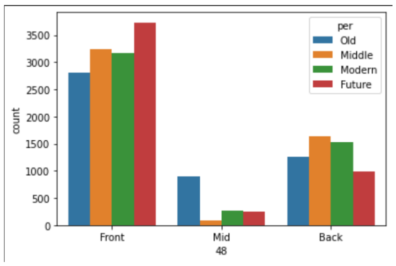
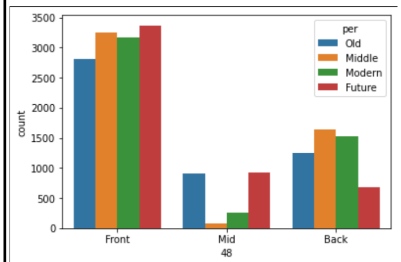
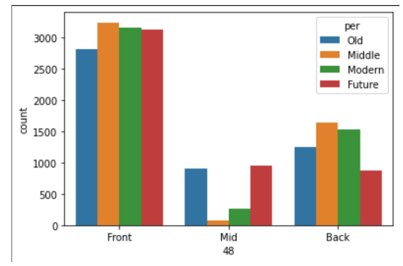
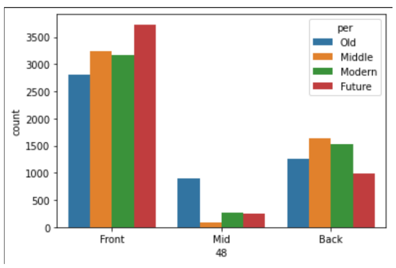
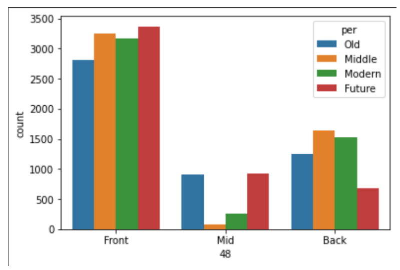
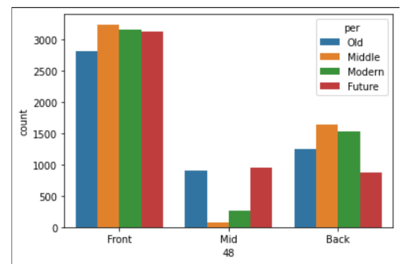
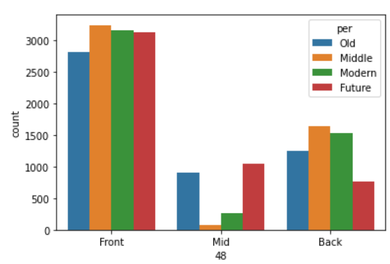
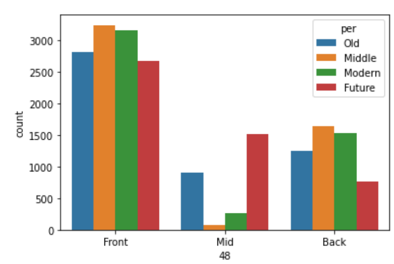
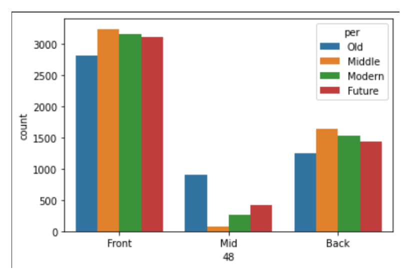
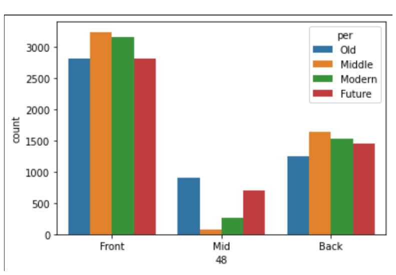
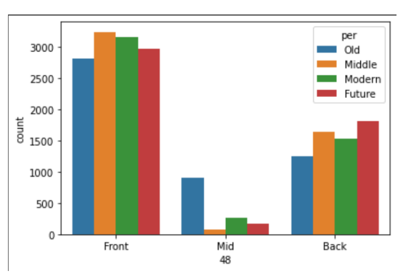
Initial consonant manner of articulation

<table>
<tr><td>Decision Tree</td><td>Random Forest</td><td>Extra Trees</td></tr>
<tr><td>
</td><td>
</td><td>
</td></tr>
<tr><td>Decision Tree MC</td><td>Random Forest MC</td><td>Extra Trees MC</td></tr>
<tr><td>
</td><td>
</td><td>
</td></tr>
<tr><td>LSTM</td><td>MLP</td><td>Polynomial 3.25</td></tr>
<tr><td>
</td><td>
</td><td>
</td></tr>
<tr><td>Polynomial 3.5</td><td>Polynomial 4</td><td></td></tr>
<tr><td>
</td><td>
</td><td></td></tr>
</table>

Initial consonant articulatory location

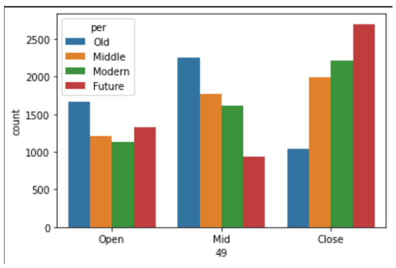
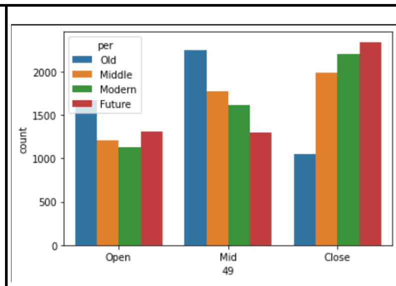
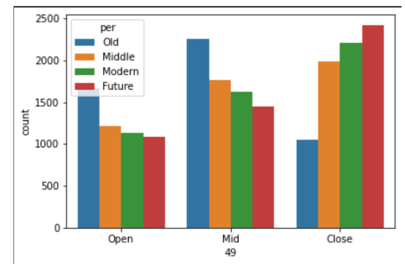
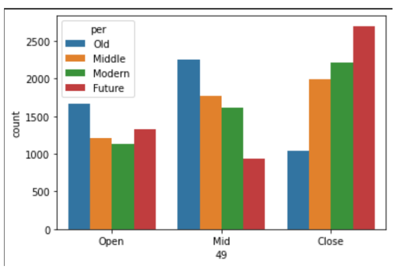
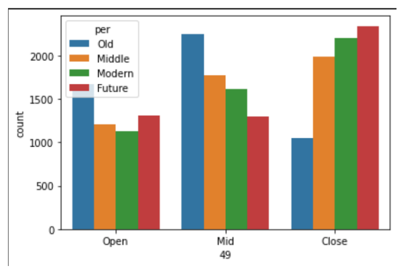
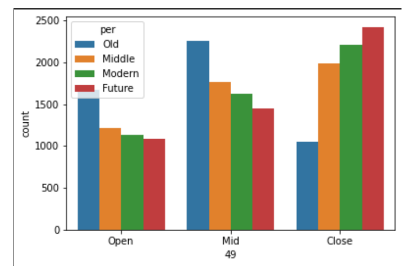
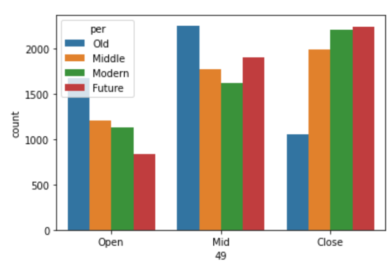
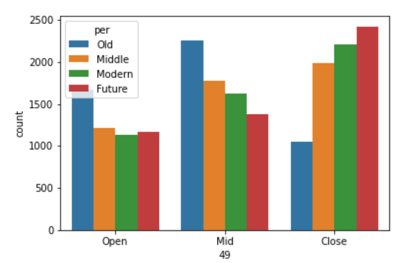
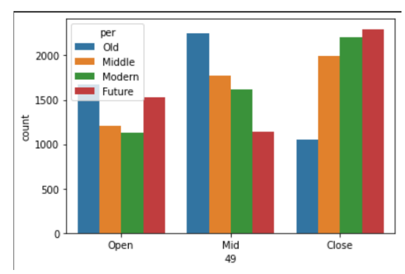
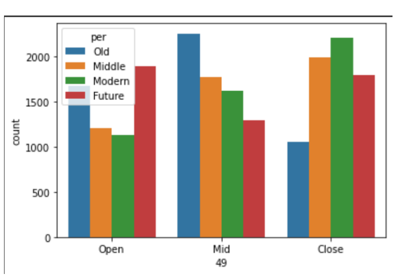
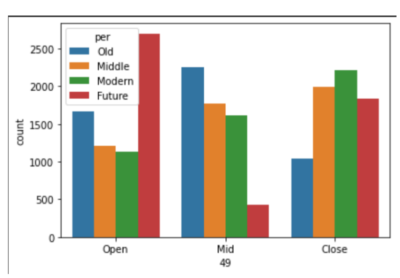
<table>
<tr><td>Decision Tree</td><td>Random Forest</td><td>Extra Trees</td></tr>
<tr>
<td>
</td>
<td>
</td>
<td>
</td>
</tr>
<tr><td>Decision Tree MC</td><td>Random Forest MC</td><td>Extra Trees MC</td></tr>
<tr>
<td>
</td>
<td>
</td>
<td>
</td>
</tr>
<tr><td>LSTM</td><td>MLP</td><td>Polynomial 3.25</td></tr>
<tr>
<td>
</td>
<td>
</td>
<td>
</td>
</tr>
<tr><td>Polynomial 3.5</td><td>Polynomial 4</td><td></td></tr>
<tr>
<td>
</td>
<td>
</td>
<td></td>
</tr>
</table>

Monophthong frontness

Decision Tree	Random Forest	Extra Trees
		
Decision Tree MC	**Random Forest MC**	**Extra Trees MC**
		
LSTM	**MLP**	**Polynomial 3.25**
		
Polynomial 3.5	**Polynomial 4**	
		

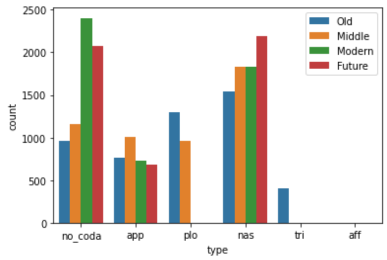
Monopthong heights

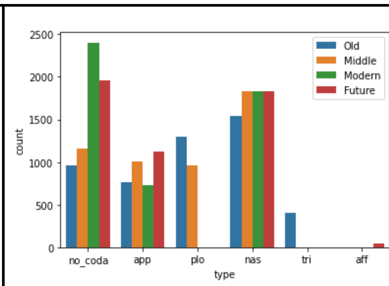
Decision Tree	Random Forest	Extra Trees

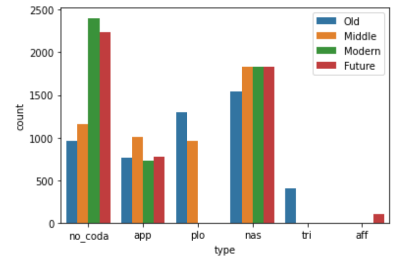
<table>
<tr><td></td><td></td><td></td></tr>
<tr><td>Decision Tree MC</td><td>Random Forest MC</td><td>Extra Trees MC</td></tr>
<tr><td></td><td></td><td></td></tr>
<tr><td>LSTM</td><td>MLP</td><td>Polynomial 3.25</td></tr>
<tr><td></td><td></td><td></td></tr>
<tr><td>Polynomial 3.5</td><td>Polynomial 4</td><td></td></tr>
<tr><td></td><td></td><td></td></tr>
</table>

Final consonant manner of articulation

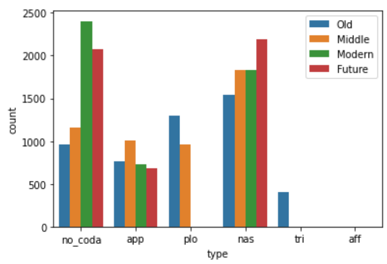
Decision Tree	Random Forest	Extra Trees

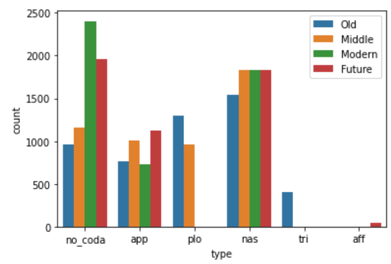


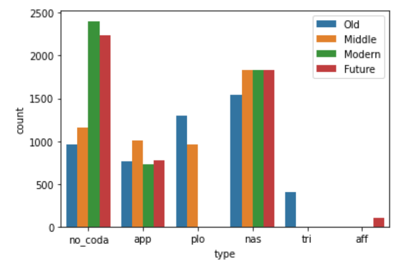




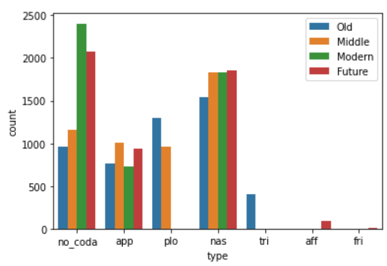
Decision Tree MC	Random Forest MC	Extra Trees MC

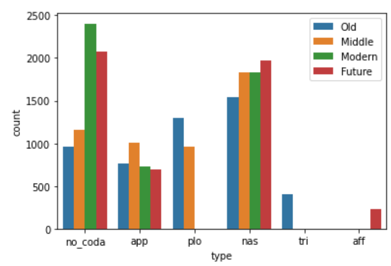


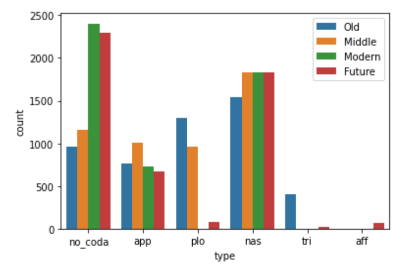




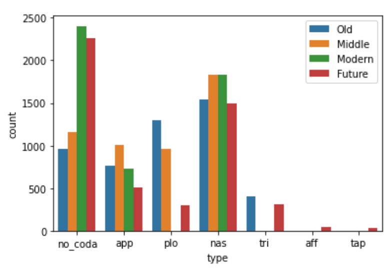
LSTM	MLP	Polynomial 3.25

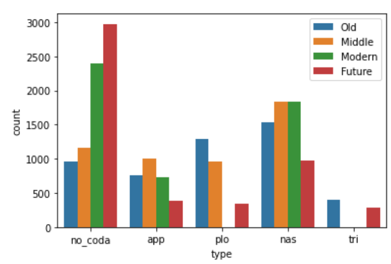






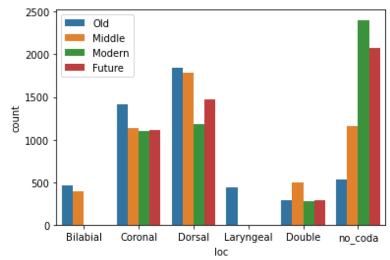
Polynomial 3.5	Polynomial 4	

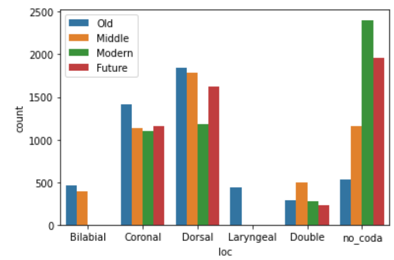


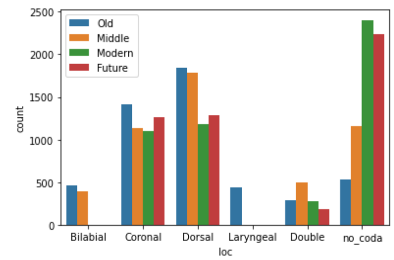


Final consonant articulatory location

Decision Tree	Random Forest	Extra Trees



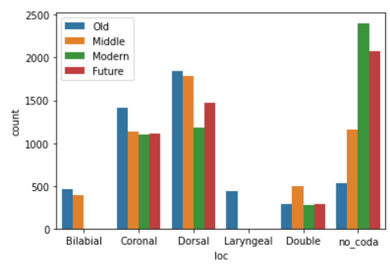


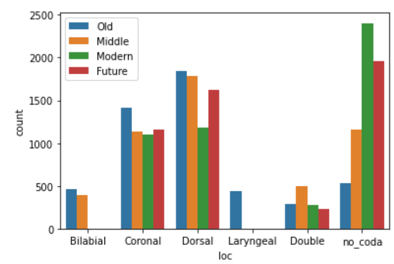


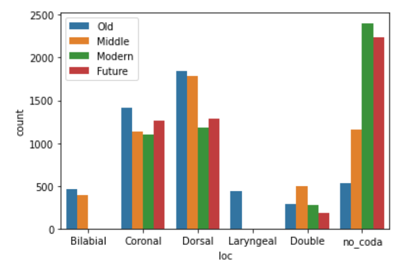
Decision Tree MC

Random Forest MC

Extra Trees MC



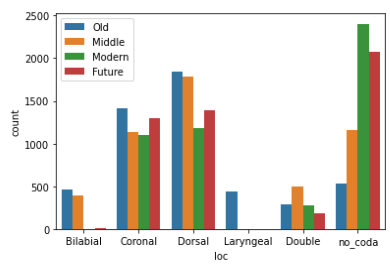


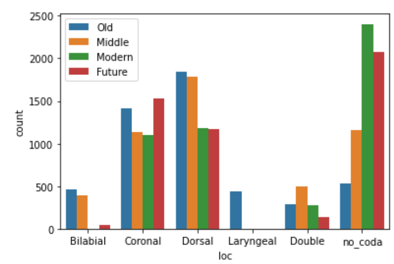


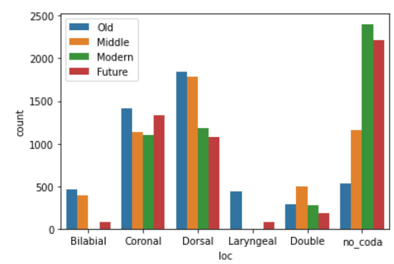
LSTM

MLP

Polynomial 3.25

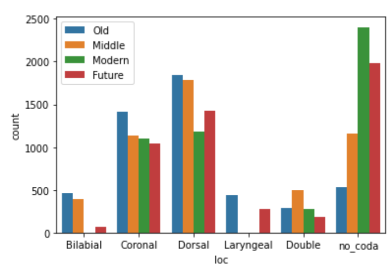


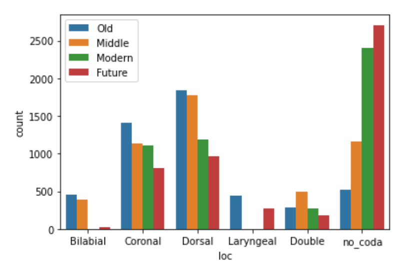




Polynomial 3.5

Polynomial 4





11. Appendix B Raw predictions of the 100 most used characters (Zhong)

Char	OC	MC_IPA	NC	Tree	Rand	Extra	Tree MC	Rand MC	Extra MC	LSTM	MLP	Poly 3.25	Poly 3.5	Poly 4
的	t-lˤewk	tek	ti	ti	θɨ	ti	ti	ti	ti	ʈʂɨ	ʈʂi	ti	c-lˤi	lˤi
一	ʔit	jit	i	i	i	i	i	i	i	i	i	i	i	ʔi
是	deʔ	dʑje	ʂʐ̩	tsɨ	tsɨ	ʃɨ	ʂɨ	ʃɨ	ʃɨ	ʈʂɘ	ɤəɰ	tʃɨ	tʃɯʔ	ɯʔ
我	ŋˤajʔ	ŋa	wɔ	pʰaj	ə	ʏ	ɕjɛ	ʃəɰ	ʂɪeɰ	y	ɵdʑ	oɰ	oj	u
不	pə	pjuw	pu	pʰu	pətɕ	pətɕ	pu	pu	pu	sɰej	θɨj	pʊ	pɘ	pæ
人	niŋ	ɲin	ʐən	tʰjɛn	ʃan	ʂɛæn	ʂan	ʃan	ʂɰɛæn	ʂəɱ	çeɱ	zɜn	ðaɳ	ɑɴ
在	dzˤəʔ	dzʌj	t͡saj	tsaj	tsaj	tsaj	tsaj	tsaj	tsaj	dʒʰəɰ	tsæadʑ	tsaj	tsˤak	tsˤa
他	l̥ˤaj	tʰa	tʰa	tʰwɔ	tʰdʑɵɔ	tsʰtɕɵə	tʰtɕɐ	tʰdʑɐ	tʰtɕɐ	dʒʰɤ	tsʰjɐ	tʰa	lˤaɰ	lˤaɹ
有	ɢʷəʔ	ɣjuw	joʊ	i	çdʑeɰ	ʏ	ʐwej	ʝdʑej	y	ɕɛɛj	χʉɴ	jɔɵ	jäɜ	ja
個	kˤar-s	ka	kɤ	tɕi	ʈʂtɕɪɰ	tɕɪ	tɕi	ʈʂtɕɪɰ	tɕɪ	tɕy	ʈʂɪ	kɯʀ-χ	kˤɯr-s	kˤɯ
上	cə-daŋʔ	dʑjaŋ	ʂaŋ	kaŋ	ʂtɕäŋ	ʂaŋ	kaŋ	ʂaŋ	ʂaŋ	ʃɰaɲ	saɳ	ʔ-tʃaŋ	c-tʃaŋ	ak
來	mə.rˤək	lʌj	laj	naj	laj	laj	laj	laj	laj	ʂɜɐɰ	læaw	laj	ɳ-rˤac	rˤak
到	tˤawk-s	taw	taw	taw	taw	taw	taw	taw	taw	ʈʂaw	tsa	taw-x	tˤaw-f	tˤaw
時	də	dʑi	ʂʐ̩	tsɨ	tsɨ	ʃɨ	ʂɨ	ʃɨ	ʃɨ	dʒʰɪ	ɘ	tʃɯ	tʃɤ	ɑ
大	lˤat-s	da	ta	tʰwɔ	tʰdʑɵɔ	tstɕɵo	twɔ	twɔ	twɔ	ʃʊ	tsjɐ	tak-x	lˤap-f	lˤa
地	lˤej-s	dij	ti	ti	tsɨ	ti	ti	ti	ti	ʈʂʰɨ	tsi	tɪ-x	lˤe-f	lˤa
為	ɢʷraj	ɣjue	wej	sɨ	ʃɪ	ʃɪɰ	tɕi	ɪ	i	zɘɵj	ɛædʑ	aɛʊ	a	a
子	tsəʔ	tsi	t͡sʐ̩	tsɨ	tsɨ	tsɨ	tsɨ	tsɨ	tsɨ	tsɨ	tsɨ	tsɯ	tsɤʔ	tsɑʔ
中	truŋ	ţjuŋ	t͡ʂʊŋ	iŋ	ʈʂɪɲ	ʈʂʊɲ	iŋ	ʈʂɪɲ	ʈʂʊɲ	ʈʂʊɲ	tsʊɲ	dʒʊŋ	tsoŋ	tsɵŋ
生	sreŋ	ʂæŋ	ʂɤŋ	ʂɤŋ	ʈʂəɲ	ʂəɲ	ʂɤŋ	ʈʂəɳ	çɪɲ	ʂəɲ	ʂʊəɲ	ʃʀɯŋ	srɯŋ	ɸɯŋ
生	sreŋ	ʂjæŋ	ʂɤŋ	ʈʂən	ʈʂəɳ	ʂəɲ	ʂɤŋ	ʈʂəɳ	çɪɲ	ʃɤŋ	ʃoɲ	ʃɯŋ	sɯŋ	ɸɯŋ

國	c.qʷˤək	kwʌk	kwɔ	wa	tɕɰəɜ	ʈʂtɕʏ	kua	cʉɛ	ctɕɵ	tɕʉ	kɵə	kwɔ	q-kˤwɘ	t-qˤwø
年	c.nˤiŋ	nen	njɛn	tɕjɛn	næn	dʑɰæn	tʰwən	nɰæn	ɹɰæn	zɰæn	ʒjɛæɱ	nɹɛn	q-nˤʊæɳ	t-nˤaɴ
著	trak-s	ʈjʌ	t͡ʂu	su	tsu	dʒʊ	tstɕo	tstɕo	ʈʂʊ	ʈʂʊ	tʃdʑɤʌ	dʒʉʔ-x	tsyc-f	y
就	dzuk-s	dzjuw	t͡ɕjoʊ	i	ʈʂdʑej	ʈʂʏ	ʂwej	ʂwej	ʈʂʊ	ʈʂɛɰ	tsdʑəɰ	tɕɰɔʊ-ç	ɔo	ɒɔ
那	nˤar	na	nwɔ	wa	tɕɛɰ	tɕʏ	ɕɵa	ɕɵɛ	ʂʏ	zu	ʐo	nɔ	nˤoɽ	nˤu
和	ɢˤoj	ɣwa	xɤ	ɕi	xtɕɪɰ	çtɕɪ	ɕi	çdʑɪɰ	çɪ	ɕyʏ	fɕəə	xɯ	kˤuɰ	qˤuɪ
要	qewk-s	jiew	jaw	aw	aw	aw	aw	aw	aw	ɕeɛj	jæaw	ʊaw-x	aw-f	aw
出	t-kʰut	tɕʰwit	t͡ʂʰu	tsu	tsʰʊ	tstɕo	tsʰtɕo	tsʰtɕo	ʈʂʰʊ	tsʰɪ	ʈʂʰɨ	ʈʂʰu	k-tʃʰu	t-u
也	lajʔ	jjæ	jɛ	sa	ʒæ	ʂʒæ	ʈʂʰaj	dʒæaɰ	ʂɰɛ	ɕɛa	jəj	ɻɛ	ɭɛɰ	lɛɪ
得	tˤək	tʌk	tɤ	tɕi	ʈʂtɕɪɰ	tsɪ	sɨ	dʒdʑəɰ	tɰɘ	ʈʂɘ	tsəj	tɯ	tˤɨ	tˤi
裏	m.rəʔ	li	li	li	li	li	li	li	li	ʈʂɨ	ʟi	lɪ	ɳ-rəʔ	rɑʔ
後	ɢˤroʔ	ɣuw	xoʊ	i	çdʑej	ʂdʑɵ	xwej	çwej	çʊ	ɕdʑej	θʊɘ	xɔo	kˤɭɒɔ	qˤɒä
自	s.bit-s	dzij	t͡sʐ	tsɨ	tsɨ	tsɨ	tsɨ	tsɨ	tsɨ	tsʰɨ	fiɘ	x-tsɯ-x	f-ɯ-f	ɯ
以	ləʔ	ji	i	i	ɪ	i	i	i	i	i	æɛ	ɪ	ɘʔ	ɑʔ
會	ɦʷˤaj-s	ɣwaj	xwej	xwej	xwej	çdʑej	xwej	çwej	xwej	ʂeɛj	ʃwej	ɦˤwij-ç	wij	wiɰ
家	kˤra	kæ	t͡ɕjä	tɕjaw	ʈʂɰɘ	ʈʂɰɘ	ʈʂwɔ	ʈʂwɔ	dʒdʑo	ʈʂʊ	ʈʂjʊɵ	tɕʊä	ɑ	ɑ
可	kʰˤajʔ	kʰa	kʰɤ	tɕʰi	ʈʂtɕɪɰ	tɕɰɪ	tɕʰɰɪ	cʰdʑeɰ	ʈʂʰɰɪ	tɕʏ	ʈʂʰɪ	kʰɯɰ	kʰˤɯj	kʰˤɯ
下	gˤraʔ	ɣæ	ɕjä	ɕyɤ	ʂɰɘə	ʂʏɘ	ʃdʑə	ʃdʑɔ	ʃdʑo	ɕɪ	ʂjɘə	ɕʊä	ɑʔ	ɑʔ
而	nə	ɲi	ɻ̩	ɘ	ɪ	ɪ	i	i	i	i	ɦɪ	ʌ	ɑ	ɑ
過	kʷˤaj	kwa	kwɔ	wa	tɕɰəɜ	ʈʂtɕʏ	kua	cʉɛ	ctɕɵ	tɕy	kɵ	kwoɰ	kˤwoj	kˤwu
天	l̥ˤin	tʰen	tʰjɛn	ʈʂan	tʰæn	tʰɰɛæn	tʰwən	tʰtɕæn	tʰɰɛæn	ʈʂʰjæn	tsʰjɛn	tʰɹɛn	lˤʊæn	lˤan
去	kʰraʔ	kʰjʌ	t͡ɕʰy	su	dʒɨ	ʈʂʰʊ	tsʰtɕo	dʒʰʊ	dʒʰtɕʊ	ʈʂʰjoʊ	ɤɑ	tɕʰy	yʔ	yʔ
能	nˤəŋ	nʌŋ	nɤŋ	niŋ	ɳɪŋ	nɪɲ	niŋ	ɳɪɲ	ɳiɲ	ʂoɲ	ɟəɳ	nɯŋ	nˤɨŋ	nˤiŋ
對	tˤup-s	twʌj	twej	kʰwej	ʈdzeɰ	ʈdzɪj	kwej	ʈdzeɰ	ʈdzej	ʈɛj	ʃej	tdʑiɪ-x	tˤy-f	tˤy

小	sewʔ	sjew	ɕjaw	aw	çadʑ	ʂtɕɛædʑ	xæw	ʂæadʑ	ʂɛædʑ	ʂwɛdʑ	ʃdʑæw	ɕɰaw	ɕɰaw	aw
多	t.lˤaj	ta	twɔ	wa	tɕəɛ	ʈʂʏ	kua	ʈʉɛ	ʈɵ	ʈʂu	ʈjʊɵ	tɔ	c-lˤoɰ	lˤuɹ
然	nan	ɲjen	ʐan	nan	san	ʂan	lan	san	ʂan	ʂɜan	çæm	zan	ðan	an
于	ɦʷa	ɣju	y	swɔ	ʏ	xtɕɵ	u	ʉ	u	ɰɵʉ	ɘɰ	y	ø	a
心	səm	sim	ɕin	ʂən	ʂən	ʂən	ʂən	ʂən	ʂən	ʂɪn	ʂin	ɕɪɳ	ɕɘɲ	ɑ
學	m-kˤruk	ɣæk	ɕyɤ	i	çtɕɪɰ	ʂɪɰ	ʂwej	ʂdʑɪɰ	ʂdʑɪɰ	ʂi	sɘ	ɕyɯ	ɳ-ʈʉu	u
之	tə	tɕi	t͡ʂʐ̩	tsɨ(tsz̩)	tsɨ(tsz̩)	dʒɨ(dʒz̩)	tsɨ(tsz̩)	tsɨ(tsz̩)	tsɨ(tsz̩)	tsɨ(tsz̩)	ʈʂɨɪ(tsz̩)	tʃɯ	tʃɤ	ɑ
都	tˤa	tu	tu	tu	tu	tu	tu	tu	tu	ʈʂʊu	tu	tʊ	tˤɜ	tˤa
好	qʰˤuʔ	xaw	xaw	aw	aw	aw	xæw	xæw	xæw	ʂjaw	ʂa	xɜdʑ-χ	qʰˤʊʔ-s	ʔʰˤu
看	kʰˤar	kʰan	kʰan	ʈʂan	cʰan	cʰan	kʰan	cʰan	kʰan	ʈʰaæn	ʈʂʰan	kʰan	kʰˤar	kʰˤar
起	c.qʰrəʔ	kʰi	t͡ɕʰi	ʈʂʰɨ	ʈʂʰɨ	ʈʂʰɨ	ʈʂʰɨ	ʈʂʰɨ	ʈʂʰɨ	ʈʂʰɨ	ʈʂʰɨ	ʔ-tɕʰʀɪ	c-rəʔ	ɑʔ
發	cə.pat	pjʌt	fa	pʰdʑɞɴ	pdʑɔ	θtɕə	pwɔ	pwɔ	sdʑɞɴ	twɔ	fwɔʌ	fa	q-fa	ʈ-sa
當	tˤaŋ	taŋ	taŋ	taŋ	taŋ	taŋ	taŋ	taŋ	taŋ	dʒaɲ	tsaŋ	taŋ	tˤaŋ	tˤaŋ
沒	mˤut	mwʌt	mwɔ	wa	ðəɛɰ	stɕɘəɰ	ɕɵa	ɕəɛ	ʃɘɰ	ndʑɔ	jdʑɔ	mwɔ	mdʑo	mˤu
成	deŋ	dʑjeŋ	t͡ʂʰɤŋ	ʈʂʰɤŋ	ʈʂʰəɲ	ʈʂɨɲ	ʈʂɤŋ	ʈʂʰəɲ	ʈʂʰɨɲ	ʈʂʰoɲ	ʈʂʰəɳ	tʃʰɤŋ	tʃʰɤŋ	ɤŋ
如	na	ɲjʌ	ʐu	nu	ʃʊ	ʃʊ	nu	su	ʂu	ɲɰʊu	xʉ	zu	ðʉ	y
事	m-s-rəʔ-s	dʐi	ʂʐ̩	tsɨ(tsz̩)	tsɨ(tsz̩)	ʃɨ(ʃz̩)	ʂɨ(ʂz̩)	ʃɨ	ʃɨ	ʈʂʰɪ	ɘ	ŋ-ʃʀʊ-χ	ɱ-srɤʔ-s	ɸɑʔ
把	pˤraʔ	pæ	pa	pwɔ	pdʑɔ	fɰʊo	pwɔ	pwɔ	pwɔ	pɵ	ptɕɐ	pa	pˤɽaʔ	pˤaʔ
還	ɢʷˤren	ɣwæn	xuan	ɕyan	χøæn	çeæn	ɕyæn	çøan	çeæn	ʂɛæn	ɕyan	xuan	ɕuæn	uɛn
用	loŋ-s	jʌŋ	jʊŋ	iŋ	ɪŋ	ʈʂʊɲ	iŋ	iɲ	ʐʊɲ	ɣʉɲ	ʊɴ	ɹuŋ-χ	luŋ-s	ʉŋ
第	lˤəjʔ-s	dej	ti	ti	tsɨ	ti	ti	ti	ti	ʈʂʰɨ	tsi	ti-x	lˤɨ-f	lˤɯ
道	lˤuʔ-s	daw	taw	taw	tʰaw	taw	taw	taw	taw	ʈʂʰa	tsa	tədʑ-x	lˤu-f	lˤu
想	saŋʔ	sjaŋ	ɕjaŋ	ʂaŋ	ʂtɕäŋ	ʂdʑäŋ	ʂaŋ	ʂaŋ	ʂaŋ	ʂwäɲ	ʃɟɐäɲ	ɕɰaŋ	ɕaŋ	ak
作	tsˤak	tsak	t͡swɔ	wa	dʒɛɰ	dʒʏ	ɕɵa	tɕɵɛ	ʈʂʉʏ	tɕy	tɕjɞɴ	tsɔ	tsˤo	tsˤu

種	k.toŋʔ	tɕjʌŋ	t͡ʂʊŋ	iŋ	ʈʂɪɲ	ʈʂʊɲ	iŋ	ʈʂɪɲ	ʈʂʊɲ	ʈʂoɲ	tsʊəɲ	ʔ-tʃuŋ	k-tʃuŋ	p-ʉk
開	kʰˤəj	kʰʌj	kʰaj	kaj	cʰaj	caj	kaj	kaj	kaj	ʈʰaj	ʈʂʰæadʑ	kʰaj	kʰˤaj	kʰˤaj
美	mrəjʔ	mij	mej	mi	mi	mi	mi	mi	mi	mi	ɱɪj	mɛj	mɽäj	mɑk
總	tsˤoŋ	tsuŋ	t͡sʊŋ	iŋ	ʈʂɪɲ	tsʊɲ	iŋ	ʈʂɪɲ	tsʊŋ	tsʊɲ	tsʊɲ	tsoŋ	tsˤɔŋ	tsˤäŋ
從	dzoŋ	dzjʌŋ	t͡sʰʊŋ	iŋ	ʈʂɪɲ	tsʰʊɲ	iŋ	ʈʂɪɲ	tsʰʊŋ	dʒʰoʊɲ	tsʰɐʊɤɲ	tsʰuŋ	tsʰuŋ	tsʰʉŋ
無	ma	mju	u	ʐu	xtɕʊ	ʂtɕo	u	u	u	ʈʂɵɕ	xɵ	ʊ	ɜ	a
情	dzeŋ	dzjeŋ	t͡ɕʰiŋ	ʈʂɤŋ	ʈʂɤɲ	ʈʂɤɲ	ʈʂʰɤŋ	ʈʂʰɤŋ	ʈʂʰɤŋ	ʈʂʰoʊɲ	ʈʂʰjɔəɳ	tɕʰiŋ	tɕʰiŋ	iŋ
己	krəʔ	ki	t͡ɕi	ʈʂɨ	ʈʂɨ	ʈʂɨ	ʈʂɨ	ʈʂɨ	ʈʂɨ	ʈʂɨ	ʈʂɨ	tɕɪ	ɽəʔ	ɑʔ
面	c.men-s	mjien	mjɛn	pɰæn	ɱæn	dzæan	pɰæn	wan	wæn	ɱjɛæn	ɱjæɱ	q-mɰaæn -x	ʈ-man-f	man
最	tsˤot-s	tswaj	t͡swej	xwej	cdʑeɰ	tɕɰeɴ	kwej	ʈʂdʑeɰ	ʈʂdʑɪj	ʈʂej	tsej	tsdʑɨɭ-x	tsˤut-f	tsˤu
女	nraʔ	ɲjʌ	ny	nu	nɨ	ɳʊ	nu	nʉ	nu	njoʊ	ɪɰ	ɱy	myʔ	yʔ
但	dˤanʔ	dan	tan	tan	tan	tan	tan	tan	tan	ʈʰan	tsan	tan	tan	tan
現	n-kˤen-s	ɣen	ɕjɛn	tɕjɛn	ʃan	ʂɰɛæn	ʂan	ʃan	ʃɰæn	ɕjen	ɕjɛn	ɴ-ɕɹæn-χ	n-ʊan-s	an
前	dzˤen	dzen	t͡ɕʰjɛn	tɕjɛn	ʈʂan	dʒɛæn	ʈʂan	ʈʂan	ʈʂæn	tɕʰɪn	ʈʂʰjɐn	tɕʰɹæn	tɕʰˤʊan	an
所	qʰaʔ	xu	swɔ	wa	tɕə	ʂʉʏ	ɕɵa	ɕɵɛ	ʂʏ	ʂʉu	ʂʉ	fä	aʔ	aʔ
同	lˤoŋ	duŋ	tʰʊŋ	iŋ	ʈɪŋ	tʰʊɲ	iŋ	ɪŋ	tʰʊŋ	ʈʂʰʊɲ	tʰɯʉɳ	tʰoŋ	lʰˤɔŋ	lʰˤäŋ
日	c.nik	ɲit	ʐ̩	ɘ	i	i	i	i	i	i	i	ɨ	q-ɯ	ʈ-ɯ
手	n̥uʔ	ɕjuw	ʂoʊ	i	ʂdzej	ʃtɕɵ	tsʰwej	ʃweɰ	ʂʊ	ʂwɛæj	θɕiɘɰ	tʃoʊ	tʃɔo	ɒɘ
又	ɢʷəʔ-s	ɣjuw	joʊ	i	çdzeɰ	ʏ	ʐwej	ʝdzej	y	ɕɛɛj	χʉɴ	jɔɵ-χ	jäɐ-s	ja
行	gˤaŋ	ɣaŋ	xaŋ	ʂwɑŋ	xaŋ	xaŋ	kaŋ	çaŋ	xaŋ	ɕjæɲ	ʂæaɳ	xaŋ	kˤaŋ	kˤaŋ
意	ʔrək-s	i	i	i	i	i	i	i	i	i	i	ʀɪʔ-x	ʔʙɜc-f	ʔɑ