



# Investigating the Predictability of an Upcoming Code-switch in Cantonese-English Bilinguals

## Background

### Phonetics of Code-switching

Research into bilingual speech has revealed a unique set of phonetic features in code-switched utterances.

Code-switching is the changing of language or dialect within an utterance.

Specifically, when approaching a code-switch there has been evidence of:

- Slowing of speech rate<sup>12</sup>
- Convergence of VOT<sup>13 14</sup>
- Convergence of vowel qualities<sup>5</sup>
- Lowering of f0<sup>6</sup>

Additionally, bilingual listeners are can utilize these phonetic cues to then facilitate the processing of an upcoming code-switch.<sup>17</sup>

This potentially implies that **upcoming code-switches are predictable**.

## Results

F1/F2<sup>18,19</sup>

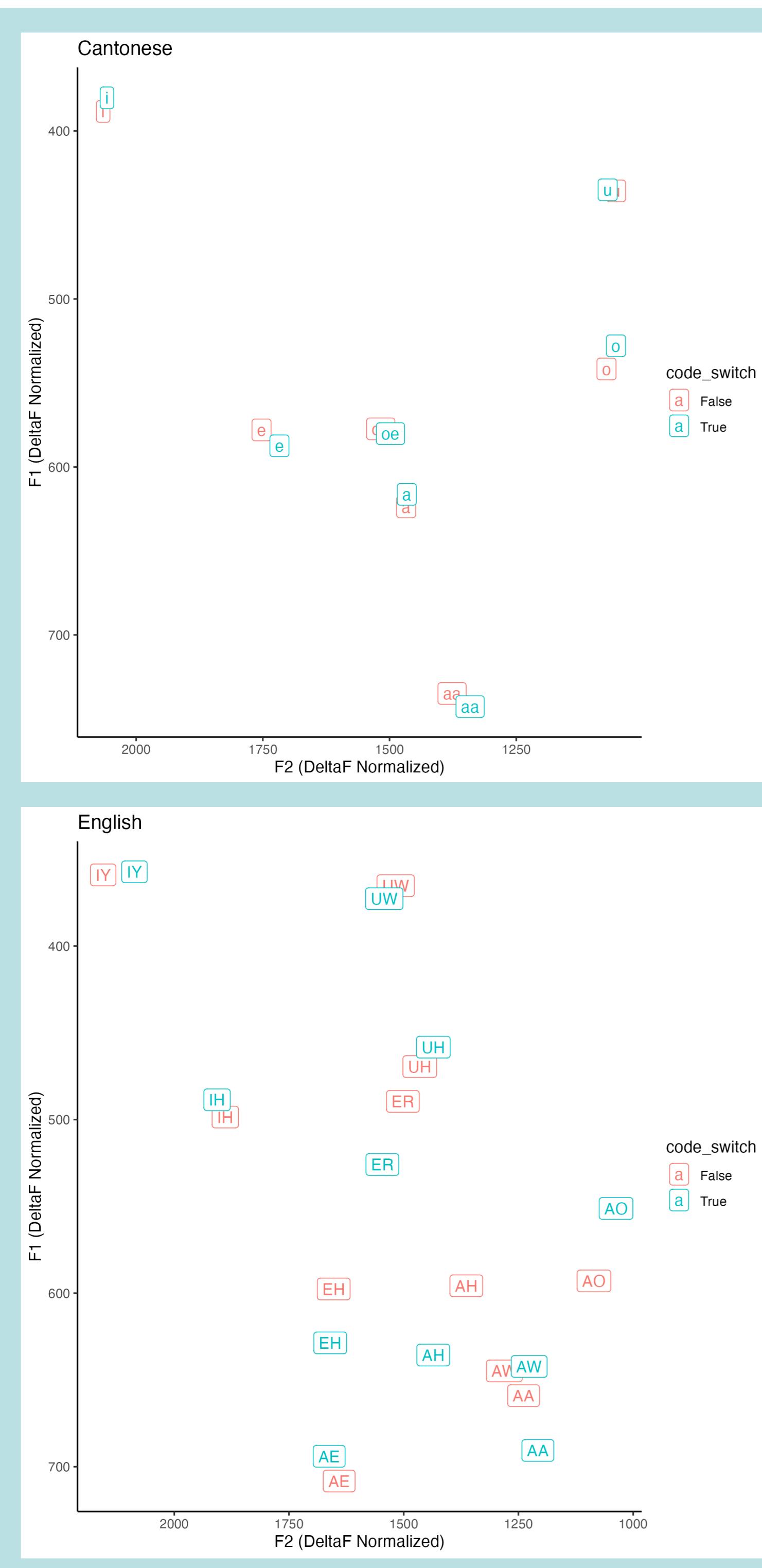
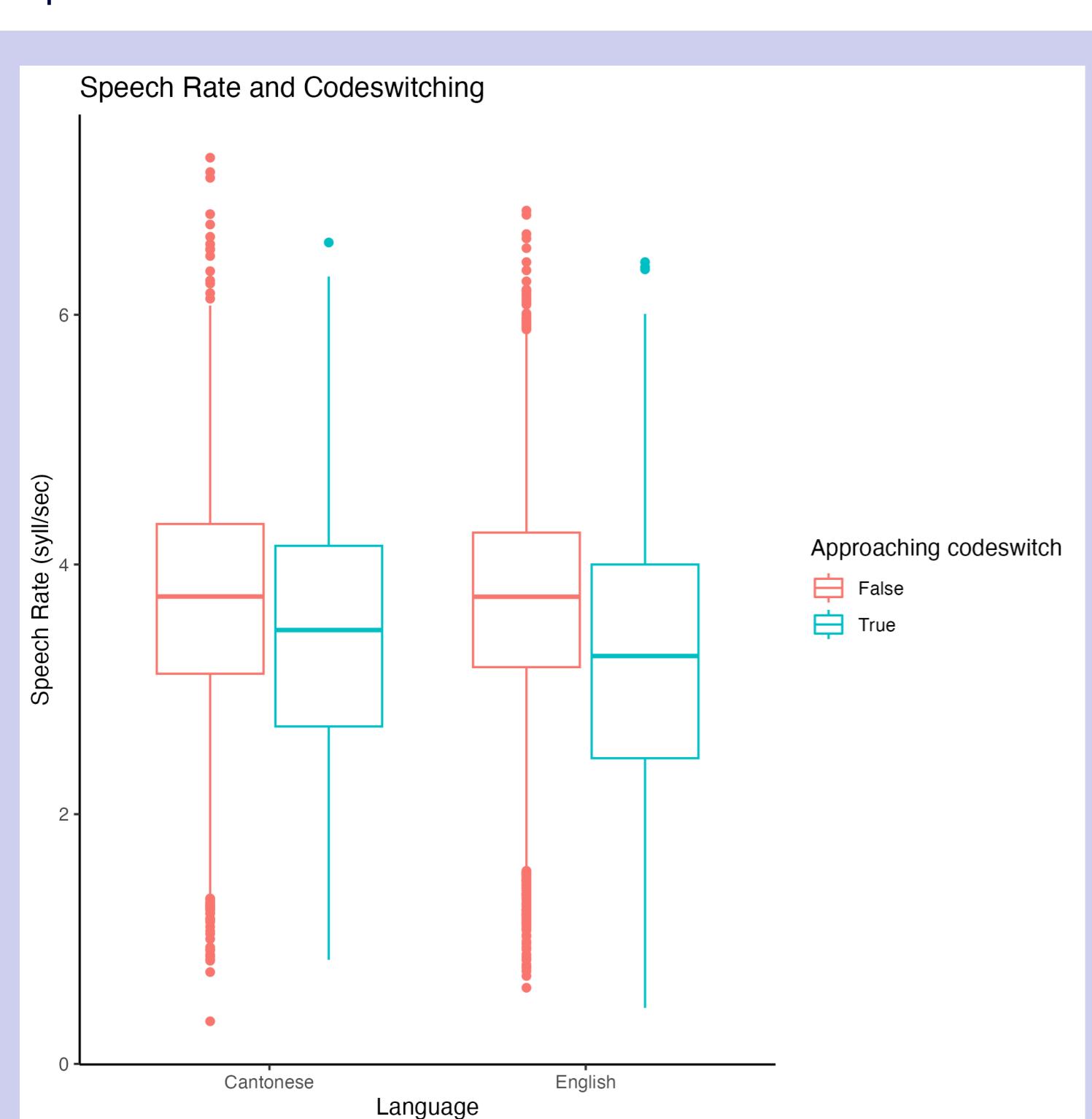


Figure 1 (Top): Vowel plot for Cantonese utterances colored by monolingual utterances (red) and code-switched utterances (blue)

Figure 2 (Bottom): Same graph for English utterances

### Speech Rate



### Neural Networks

Neural networks (NNs) are a machine learning tool used for the purpose of classification. It can be trained on a data set, then use the training to learn to classify future inputs.

NNs have been used in a variety of ways in linguistics. For example, multilingual speech segmentation<sup>8</sup>, phonetic to phonological categorization<sup>9</sup>, acoustic/phonetic feature modeling<sup>10</sup>, and phonological feature detection<sup>11</sup>

### The Present Study

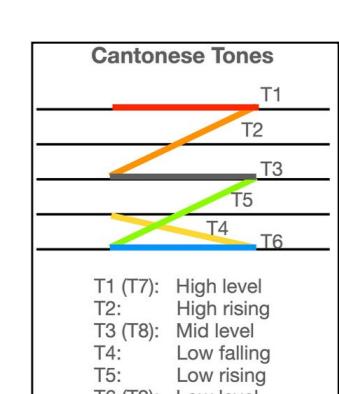
The present study investigates whether an upcoming code-switch is predictable in the SpICE corpus<sup>12</sup> of Cantonese-English Bilinguals (n = 34) in Vancouver BC using a combination of regression models and neural networks in English and Cantonese.

### Cantonese and English Phonology

Cantonese has ~10 monophthong vowel phonemes

English has ~11 monophthongs vowel phonemes

Cantonese has 6 contrastive tones<sup>6</sup>:



### Methodology

Utterances with more than one code-switch were only analyzed up to the first code-switch

Within a code-switched utterance, only the part leading up to the code-switch was analyzed

F1-4<sup>13, 14</sup>, f0<sup>15</sup> and speech rate<sup>16</sup> were analyzed.

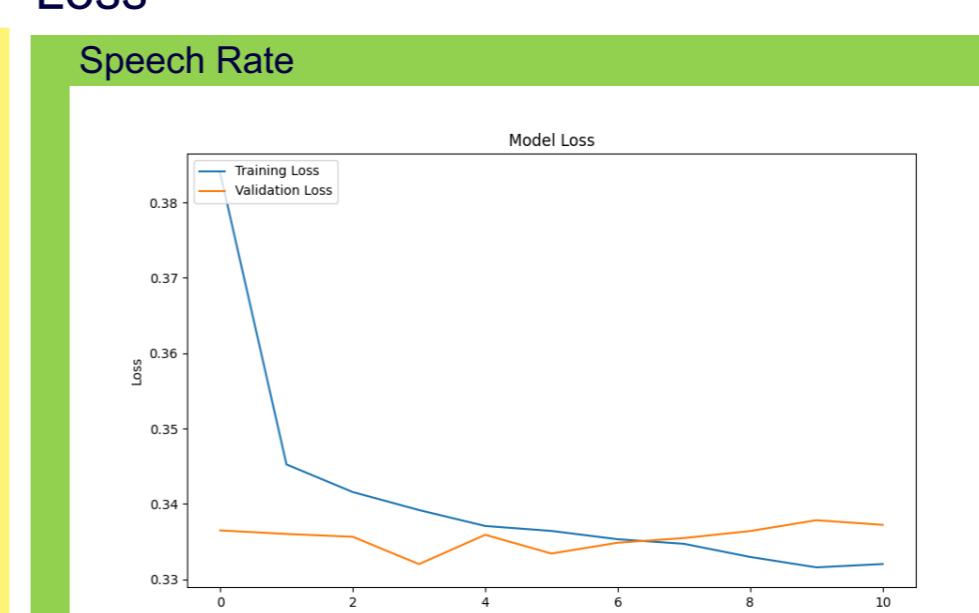
Only monophthongs were analyzed for formant analysis

Vowels that did not have a codeswitched counterpart did not get analyzed

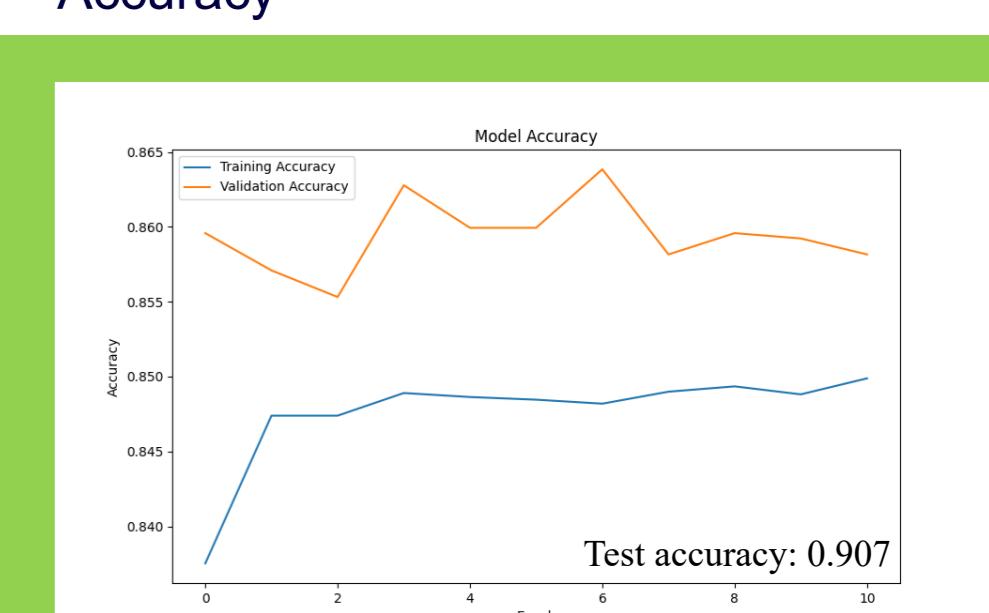
Only vowels with primary stress were analyzed

### Neural Networks

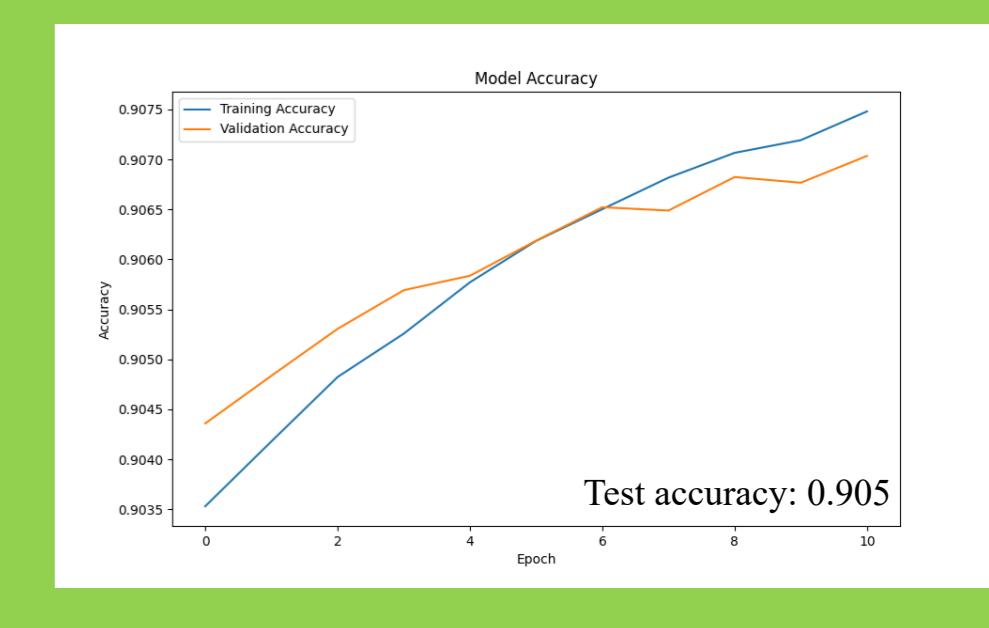
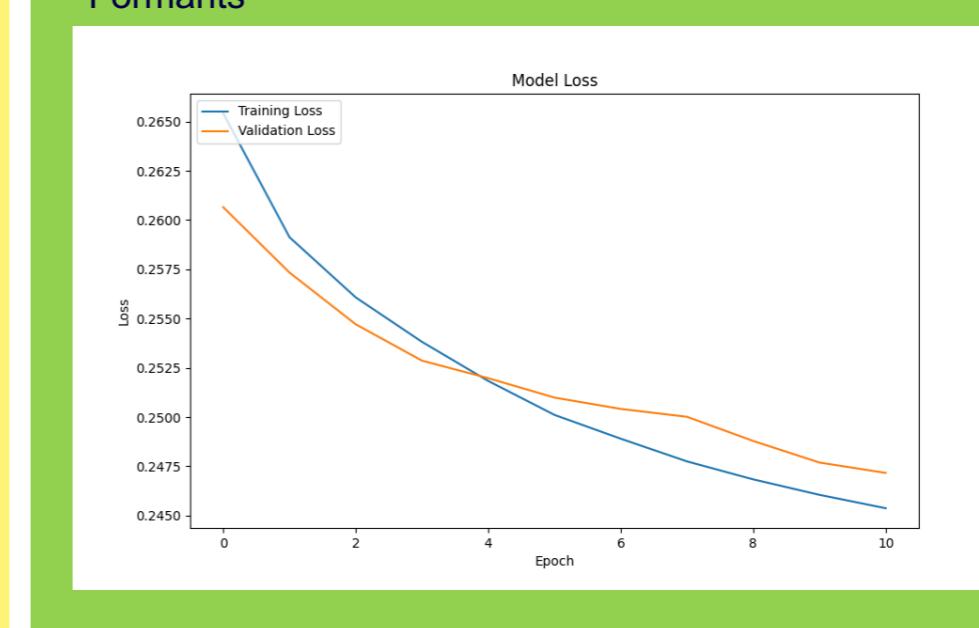
#### Loss



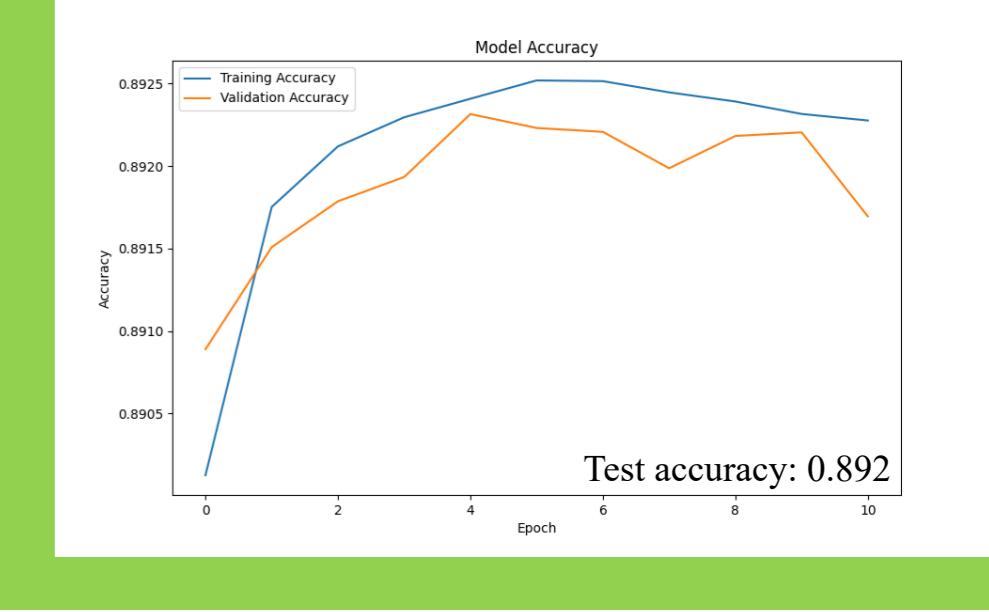
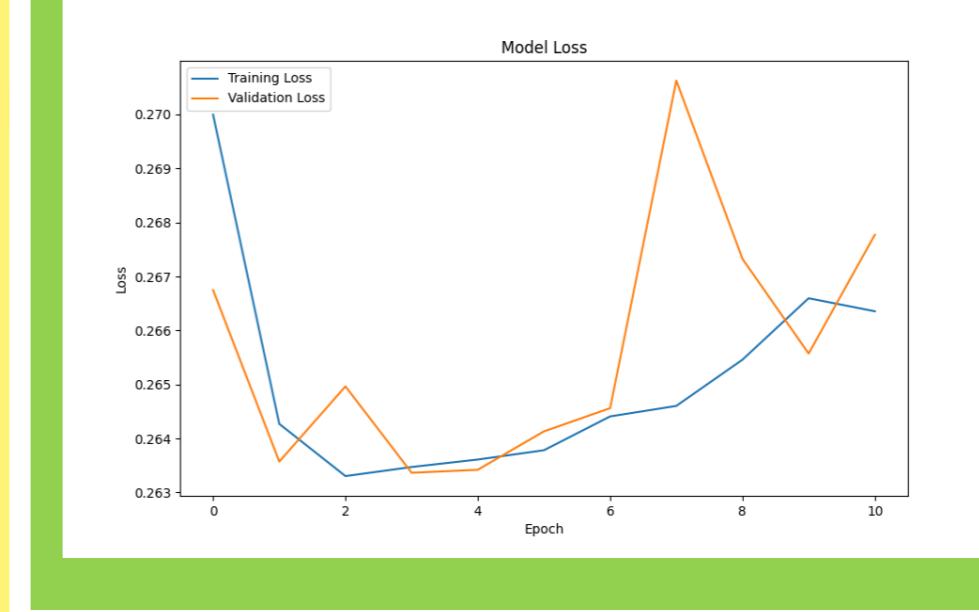
#### Accuracy



#### Formants



#### f0



### Regression Models (given codeswitch is true)

F0: f0~ tone \* codeswitch + (1|speaker)

	Estimate	P-value
tone3	2.508e+00	<0.05
tone4	1.953e+00	<0.05
tone5	1.239e+00	<0.05
English	6.505e+00	<0.05

F2: f2~ phone \* codeswitch + (1|speaker)

	Estimate	P-value
aa	-3.690e+01	<0.05
ee	-3.126e+01	<0.05
i	-7.552e+00	0.02403
o	-1.605e+01	<0.05
oe	-3.309e+01	<0.05
u	1.028e+01	0.00757

### English

	Estimate	P-value
AA	-4.964e+01	0.00105
AO	-5.164e+01	<0.005
ER	7.531e+01	<0.005
IH	-3.128e+01	<0.05
IY	-7.383e+01	0.00177

### Speech Rate

	Estimate	P-value
Intercept	3.732e+00	<0.05
English	-8.011e-02	<0.05
csTRUE	-2.433e-01	<0.05
English:csTrue	-1.464e-01	<0.05

## References



### Conclusions

The results of this study show evidence of upcoming code-switch cues in vowel formants, f0, and speech rate

Specifically, an upcoming code switch cause Cantonese /a/ and /e/ and English /a/ to lower before a code-switch, and English /æ/ and /ɔ/ to raise. Additionally, Cantonese /a/, /e/, /i/, /l/, /o/ and /œ/, and English /a/, /ɔ/, /i/ and /l/ to become more back while Cantonese /u/ and English /ə/ become more fronted before a codeswitch.

Tones 3, 4, and 5 were found to be significantly higher in f0 before a codeswitch, and there was a significant raise in f0 in English before a codeswitch.

Speech rate shows significant decrease before a code-switch, but even more so when code-switching from English.

The implementation of Neural Networks shows some success. The three neural networks found around 90% accuracy for test data after training. Speech rate did slightly better than the other two models, but not by much.