## Supplementary Materials Code-Switching Language Modeling using Syntax-Aware Multi-Task Learning

Genta Indra Winata, Andrea Madotto, Chien-Sheng Wu, Pascale Fung

## 1 Results

We show the performance of each model with its hyper-parameter setting.

Table 1: Results in Phase I

| Model                     | Hidden | Embedding | Dropout | POS tag dropout | PPL    | PPL    |
|---------------------------|--------|-----------|---------|-----------------|--------|--------|
| Model                     | size   | size      |         |                 | dev    | test   |
| LSTM                      | 200    | 200       | 0.2     | -               | 197.5  | 196.84 |
| LSTW                      | 500    | 500       | 0.4     | -               | 190.33 | 185.91 |
| + syntactic features      | 200    | 200       | 0.2     | -               | 187.37 | 184.87 |
| + syntactic reatures      | 500    | 500       | 0.4     | -               | 178.51 | 176.57 |
| Multi-task ( $p = 0.25$ ) | 200    | 200       | 0.4     | 0.2             | 180.91 | 178.18 |
| With-task (p = 0.20)      | 500    | 500       | 0.4     | 0.4             | 173.55 | 174.96 |
| Multi-task ( $p = 0.50$ ) | 200    | 200       | 0.4     | 0.2             | 182.6  | 178.75 |
| (p = 0.50)                | 500    | 500       | 0.4     | 0.4             | 175.23 | 173.89 |
| Multi-task ( $p = 0.75$ ) | 200    | 200       | 0.4     | 0.2             | 180.90 | 178.18 |
| p = 0.15                  | 500    | 500       | 0.4     | 0.4             | 185.83 | 178.49 |

Table 2: Results in Phase II

| Model                     | Hidden | Embedding    | Dropout         | POS tag dropout | PPL    | PPL    |
|---------------------------|--------|--------------|-----------------|-----------------|--------|--------|
| Model                     | size   | size Dropout | FOS tag dropout | dev             | test   |        |
| RNNLM                     | 200    | 200          | -               | -               | 181.87 | 176.80 |
| KININLIVI                 | 500    | 500          | -               | -               | 178.35 | 171.27 |
| LSTM                      | 200    | 200          | 0.2             | -               | 156.77 | 159.58 |
| LSTM                      | 500    | 500          | 0.4             | -               | 150.65 | 153.06 |
| + syntactic features      | 200    | 200          | 0.2             | -               | 153.6  | 152.66 |
|                           | 500    | 500          | 0.4             | -               | 147.44 | 148.38 |
| Multi-task ( $p = 0.25$ ) | 200    | 200          | 0.4             | 0.2             | 149.68 | 149.84 |
| (p - 0.25)                | 500    | 500          | 0.4             | 0.4             | 141.86 | 141.71 |
| Multi-task ( $p = 0.50$ ) | 200    | 200          | 0.4             | 0.2             | 150.92 | 152.38 |
| (p - 0.50)                | 500    | 500          | 0.4             | 0.4             | 144.18 | 144.27 |
| Multi-task ( $p = 0.75$ ) | 200    | 200          | 0.4             | 0.2             | 150.32 | 151.22 |
| p = 0.75                  | 500    | 500          | 0.4             | 0.4             | 145.08 | 144.85 |

## 2 Recording Lists

We split the recording ids into train, development, and test set as the following:

Table 3: Recording distribution in Phase I

| Data  | Recording list                  |                         |  |  |  |
|-------|---------------------------------|-------------------------|--|--|--|
| Data  | Conversation                    | Interview               |  |  |  |
|       |                                 | NI02FAX,NI04FBX,NI05MBQ |  |  |  |
|       |                                 | NI06FBP,NI07FBQ,NI08FBP |  |  |  |
|       |                                 | NI09FBP,NI10FBP,NI11FBP |  |  |  |
|       |                                 | NI12MAP,NI13MBQ,NI14MBP |  |  |  |
|       | 02NC03FBX, 02NC04FBY, 03NC05FAX | NI15FBQ,NI16FBP,NI17FBQ |  |  |  |
|       | 03NC06FAY,04NC07FBX,04NC08FBY   | NI18MBP,NI19MBQ,NI20MBP |  |  |  |
|       | 06NC11MAX,06NC12MAY,07NC13MBP   | NI21MBQ,NI22FBP,NI23FBQ |  |  |  |
|       | 07NC14FBQ,08NC15MBP,08NC16FBQ   | NI24MBP,NI25MBQ,NI26FBP |  |  |  |
|       | 09NC17FBP,09NC18MBQ,10NC19MBP   | NI27MBQ,NI28MBP,NI29MBP |  |  |  |
|       | 10NC20MBQ,11NC21FBP,11NC22MBQ   | NI30MBQ,NI31FBP,NI32FBQ |  |  |  |
|       | 12NC23FBP,12NC24FBQ,13NC25MBP   | NI33MBP,NI34FBQ,NI35FBP |  |  |  |
|       | 13NC26MBQ,14NC27MBP,14NC28MBQ   | NI36MBQ,NI37MBP,NI39FBP |  |  |  |
|       | 16NC31FBP,16NC32FBQ,18NC35FBP   | NI40FBQ,NI41MBP,NI42FBQ |  |  |  |
|       | 18NC36MBQ,19NC37MBP,19NC38FBQ   | NI43FBP,NI44MBQ,NI45FBP |  |  |  |
|       | 21NC41MBP,22NC43FBP,22NC44MBQ   | NI46FBQ,NI47MBP,NI48FBQ |  |  |  |
| Train | 23NC35FBQ,23NC45MBP,24NC35FBQ   | NI49MBP,NI50FBQ,NI51MBP |  |  |  |
|       | 24NC45MBP,25NC43FBQ,25NC47MBP   | NI52MBQ,NI53FBP,NI54FBQ |  |  |  |
|       | 26NC48FBP,26NC49FBQ,27NC47MBQ   | NI55FBP,NI56MBX,NI57FBQ |  |  |  |
|       | 27NC50FBP,28NC51MBP,28NC52FBQ   | NI58FBP,NI59FBQ,NI60MBP |  |  |  |
|       | 29NC53MBP,29NC54FBQ,30NC48FBP   | NI61FBP,NI62MBQ,NI63MBP |  |  |  |
|       | 30NC49FBQ,31NC35FBQ,31NC50XFB   | NI64FBQ,NI65MBP,NI66MBQ |  |  |  |
|       | 32NC36MBQ,32NC50FBP,33NC37MBP   | NI67MBQ,UI02FAZ,UI03FAZ |  |  |  |
|       | 33NC43FBQ,34NC37MBP,35NC56MBP   | UI04FAZ,UI05MAZ,UI06MAZ |  |  |  |
|       | 36NC46FBQ,37NC45MBP,38NC50FBP   | UI07FAZ,UI08MAZ,UI10FAZ |  |  |  |
|       | 39NC57FBX,40NC58FAY,41NC59MAX   | UI11FAZ,UI12FAZ,UI13FAZ |  |  |  |
|       | 42NC60FBQ,44NC44MBQ,45NC22MBQ   | UI14MAZ,UI15FAZ,UI16MAZ |  |  |  |
|       | 46NC41MBP,46NC41MBP             | UI17FAZ,UI18MAZ,UI19MAZ |  |  |  |
|       |                                 | UI20MAZ,UI21MAZ,UI22MAZ |  |  |  |
|       |                                 | UI23FAZ,UI24MAZ,UI25FAZ |  |  |  |
|       |                                 | UI26MAZ,UI27FAZ,UI28FAZ |  |  |  |
|       |                                 | UI29FAZ                 |  |  |  |
| Dev   | 01NC01FBX, 01NC02FBY, 15NC29FBP | UI01FAZ, UI09MAZ        |  |  |  |
|       | 15NC30MBQ, 21NC42MBQ, 43NC61FBQ | 31011712, 310711112     |  |  |  |
| Test  | 05NC09FAX, 05NC10MAY, 17NC33FBP | NI01MAX, NI03FBX        |  |  |  |
| 1631  | 17NC34FBQ, 20NC39MBP, 20NC40FBQ |                         |  |  |  |

Table 4: Recording distribution in Phase II

| Data  | Recording distribution in Phase II  Recording list  |   |  |  |  |
|-------|---|---|--|--|--|
| Data  | Conversation  | Interview   |  |  |  |
| Train | 02NC03FBX, 02NC04FBY, 03NC05FAX<br>03NC06FAY, 04NC07FBX, 04NC08FBY<br>06NC11MAX, 06NC12MAY, 07NC13MBP<br>07NC14FBQ, 08NC15MBP, 08NC16FBQ<br>09NC17FBP, 09NC18MBQ, 10NC19MBP<br>10NC20MBQ, 11NC21FBP, 11NC22MBQ<br>12NC23FBP, 12NC24FBQ, 13NC25MBP<br>13NC26MBQ, 14NC27MBP, 14NC28MBQ<br>16NC31FBP, 16NC32FBQ, 18NC35FBP<br>18NC36MBQ, 19NC37MBP, 19NC38FBQ<br>21NC41MBP, 22NC43FBP, 22NC44MBQ<br>23NC35FBQ, 23NC45MBP, 24NC35FBQ<br>24NC45MBP, 25NC43FBQ, 25NC47MBP<br>26NC48FBP, 26NC49FBQ, 27NC47MBQ<br>27NC50FBP, 28NC51MBP, 28NC52FBQ<br>29NC53MBP, 29NC54FBQ, 30NC48FBP<br>30NC49FBQ, 31NC35FBQ, 31NC50XFB<br>32NC36MBQ, 32NC50FBP, 33NC37MBP<br>33NC43FBQ, 34NC37MBP, 35NC56MBP<br>36NC46FBQ, 37NC45MBP, 38NC50FBP<br>39NC57FBX, 40NC58FAY, 41NC59MAX<br>42NC60FBQ, 44NC44MBQ, 45NC22MBQ<br>46NC41MBP | NI02FAX, NI04FBX, NI05MBQ NI06FBP, NI07FBQ, NI08FBP NI09FBP, NI10FBP, NI12MAP NI13MBQ, NI14MBP, NI15FBQ NI16FBP, NI17FBQ, NI18MBP NI19MBQ, NI20MBP, NI21MBQ NI22FBP, NI23FBQ, NI24MBP NI25MBQ, NI26FBP, NI27MBQ NI28MBP, NI29MBP, NI30MBQ NI31FBP, NI32FBQ, NI33MBP NI34FBQ, NI35FBP, NI36MBQ NI37MBP, NI39FBP, NI40FBQ NI41MBP, NI42FBQ, NI43FBP NI44MBQ, NI45FBP, NI46FBQ NI47MBP, NI48FBQ, NI55FBP NI50FBQ, NI51MBP, NI52MBQ NI53FBP, NI54FBQ, NI55FBP NI56MBX, NI57FBQ, NI58FBP NI59FBQ, NI60MBP, NI61FBP NI62MBQ, NI63MBP, NI64FBQ NI65MBP, NI66MBQ, NI67MBQ UI02FAZ, UI03FAZ, UI04FAZ UI05MAZ, UI06MAZ, UI07FAZ UI08MAZ, UI10FAZ, UI11FAZ UI12FAZ, UI13FAZ, UI14MAZ UI15FAZ, UI16MAZ, UI17FAZ UI18MAZ, UI19MAZ, UI20MAZ UI21MAZ, UI22MAZ, UI23FAZ UI24MAZ, UI25FAZ, UI26MAZ UI27FAZ, UI28FAZ, UI29FAZ |  |  |  |
| Dev   | 15NC30MBQ, 21NC42MBQ, 43NC61FBQ   | UI01FAZ, UI09MAZ  |  |  |  |
| Test  | 05NC09FAX, 05NC10MAY, 17NC33FBP<br>17NC34FBQ, 20NC39MBP, 20NC40FBQ  | NI01MAX, NI03FBX  |  |  |  |