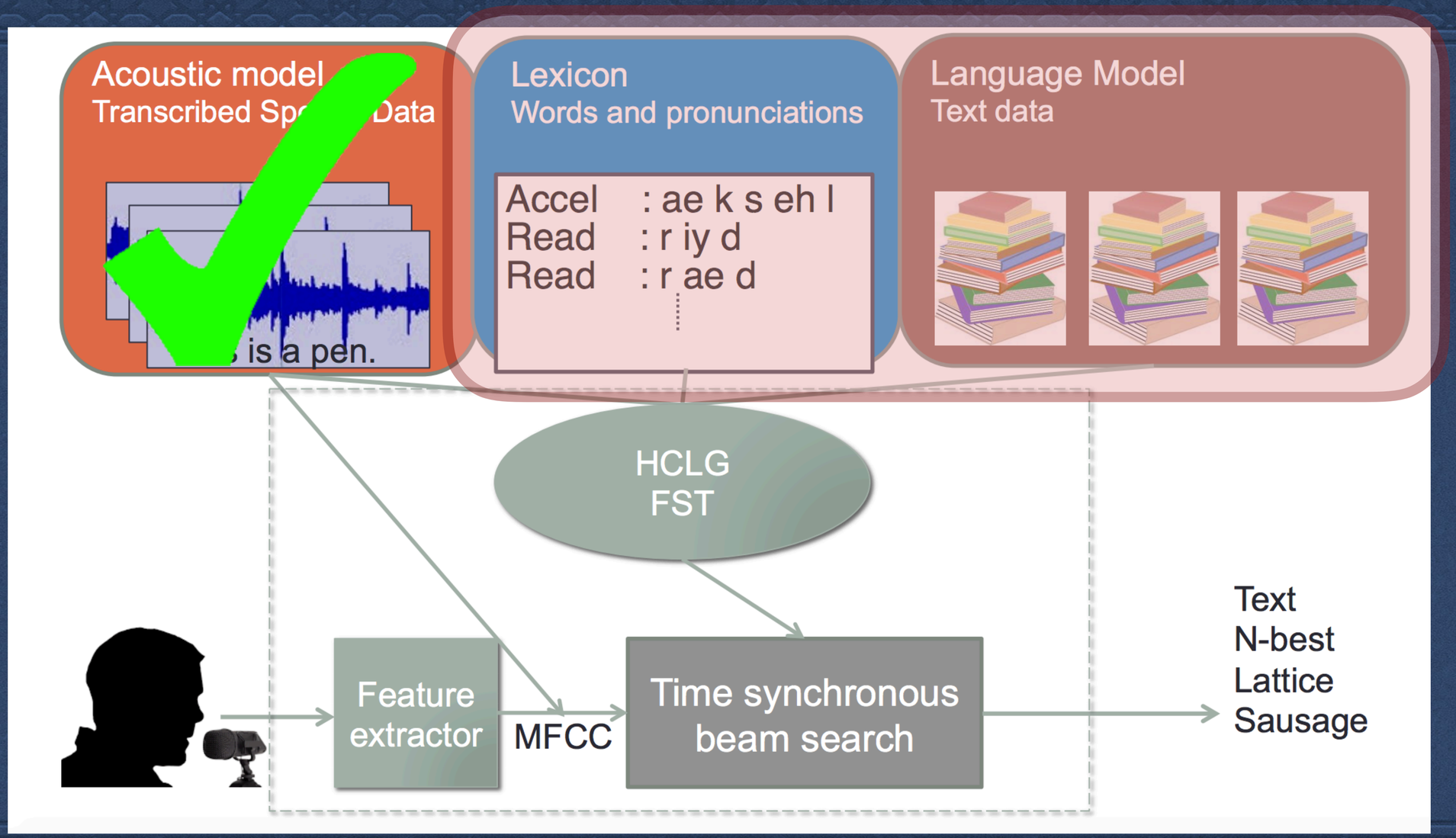
กระหึ่มเพลเยอร์

2110432 Auto Speech Recognition

What we tried ?


Acoustic model
Transcribed Sp Data
is a pen.
Lexicon
Words and pronunciations
Accel : ae k s eh l
Read : r iy d
Read : r ae d
Language Model
Text data
HCLG
FST
Feature
extractor
MFCC
Time synchronous
beam search
Text
N-best
Lattice
Sausage

AM.tar.gz

```
exp
└── nnet2_online
    ├── conf
    │   ├── ivector_extractor.conf
    │   ├── mfcc.conf
    │   ├── online_cmvn.conf
    │   ├── online_nnet2_decoding.conf
    │   └── splice.conf
    ├── final.mdl
    ├── ivector_extractor
    │   ├── final.dubm
    │   ├── final.ie
    │   ├── final.mat
    │   ├── global_cmvn.stats
    │   ├── num_jobs
    │   ├── online_cmvn.conf
    │   └── splice_opts
    ├── phones.txt
    └── tree
```

Lexicon

Words and pronunciations

Accel	: ae k s eh l
Read	: r iy d
Read	: r ae d
	⋮

Train ข้อมูล
และใช้ Sequitur G2P

testlex.txt

เล่น

หยุด

เปลี่ยน

เพลง

ไกลแค่ไหนคือใกล้

คู่กัน

ชะตาชีวิต

กะทันหัน

ฝน

g2plex.txt

```
เล่น0    0.997698 l e n^
หยุด     0    0.998954 j u t^
เปลี่ยน  0    0.999723 pl iia n^
เพลง     0    0.990766 phl ee ng^
ไกลแค่ไหนคือใกล้   0    0.167739 kl a j^ kh xx n a kh vv kl a j^
คู่กัน   0    0.896848 kh uu k a n^
ชะตาชีวิต   0    0.992171 ch a t aa ch ii w i t^
กะทันหัน    0    0.999510 k a th a n^ h a n^
ฝน 0    0.971064 f o n^
```

Language Model

Text data

สร้าง speech recognizer จาก lexicon กับ LM

คำสั่งที่ใช้ในแอปพลิเคชัน

กริยา		กรรม
เล่น		เพลง
หยุด		ไกลแค่ไหนคือใกล้
เปลี่ยน	+	คู่กัน
		ชะตาชีวิต
		กะทันหัน
		ฝน

เล่น
หยุด
เปลี่ยน
เพลง
คู่กัน
ชะตาชีวิต
กะทันหัน
ฝน
ไกลแค่ไหนคือใกล้

1. Creating L and G fst

mylexicon.txt

เล่น	l e n^
หยุด	j u t^
เปลี่ยน	pl iia n^
เพลง	phl ee ng^
ไกลแค่ไหนคือใกล้	kl a j^ kh xx n a kh vv kl a j^
คู่กัน	kh uu k a n^
ชะตาชีวิต	ch a t aa ch ii w i t^
กะทันหัน	k a th a n^ h a n^
ฝน	f o n^

Generated using SRILM (mylm.arpa)

```
\data\
ngram 1=11
ngram 2=25
ngram 3=14
ngram 4=6
```

mylm.arpa

```
\1-grams:
-0.5827549    </s>
-99    <s>-1.20658
-1.342423    กะทันหัน-0.5674007
-1.342423    คู่กัน    -0.5674007
-1.342423    ชะตาชีวิต    -0.5674007
-1.342423    ฝน  -0.5674007
-0.8653014    หยุด    -0.6879748
-1.944483    เปลี่ยน   -0.2138798
-0.7403627    เพลง    -0.9391787
-0.7403627    เล่น -0.7483641
-1.342423    ไกลแค่ไหนคือใกล้ -0.5674007
```

mylm.arpa

```
\2-grams:
-0.30103  <s> หยุด      0.4648868
-1.380211      <s> เปลี่ยน
-0.3802112    <s> เล่น 0.3577674
-0.09691001  กะทันหัน </s>
-0.09691001  คู่กัน </s>
-0.09691001  ชะตาชีวิต </s>
-0.09691001  ฝน </s>
-1.113943      หยุด </s>
-0.4149733    หยุด เพลง
-0.3357921    หยุด เล่น      0.3383543
.
.
.
```

mylm.arpa

```
\3-grams:
-0.1249387    เพลง กะทันหัน </s>
-0.1249387    เพลง คู่กัน </s>
-0.1249387    เพลง ชะตาชีวิต </s>
-0.1249387    เพลง ฝน </s>
-1.290035     <s> หยุด เพลง
-0.30103  <s> หยุด เล่น      -0.7311547
-0.7403627    เล่น เพลง กะทันหัน
0.1249388
-0.7403627    เล่น เพลง คู่กัน    0.1249388
-0.7403627    เล่น เพลง ชะตาชีวิต
                    .
                    .
                    .
```

mylm.arpa

```
\4-grams:
-0.1760913    เล่น เพลง กะทันหัน </s>
-0.1760913    เล่น เพลง คู่กัน </s>
-0.1760913    เล่น เพลง ชะตาชีวิต </s>
-0.1760913    เล่น เพลง ฝน </s>
-0.07918125   <s> หยุด เล่น เพลง
-0.1760913    เล่น เพลง ไกลแค่ไหนคือใกล้ </s>
```

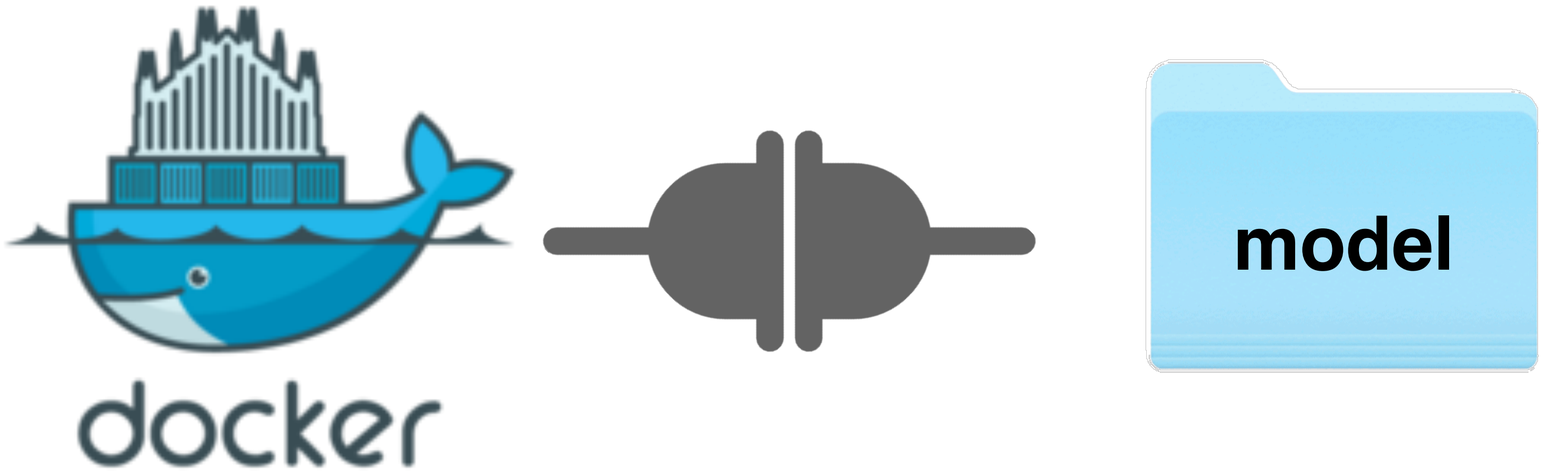
words.txt ที่ได้จากขั้นตอน

decode ของ การสร้าง HCLG

<eps> 0

!SIL 1

ฤะทันหัน 2


docker
model

Web application

(based on dictate.js)

กระหึ่มเพลเยอร์

รายการคำสั่ง

ลำดับ	ชื่อเพลง	สถานะ
1	ไกลแค่ไหนคือใกล้	หยุด
2	คู่กัน	หยุด
3	ชะตาชีวิต	หยุด
4	กะทันหัน	หยุด
5	ฝน	หยุด

เริ่มคำสั่งเสียง ส่งคำสั่งเสียง

กระหึ่มเพลเยอร์

รายการคำสั่ง

ลำดับ	ชื่อเพลง	สถานะ
1	ไกลแค่ไหนคือใกล้	หยุด
2	คู่กัน	หยุด
3	ชะตาชีวิต	หยุด
4	กะทันหัน	หยุด
5	ฝน	หยุด

เริ่มคำสั่งเสียง

ส่งคำสั่งเสียง

กระหึ่มเพลเยอร์

รายการคำสั่ง

ลำดับ	ชื่อเพลง	สถานะ
1	ไกลแค่ไหนคือใกล้	หยุด
2	คู่กัน	หยุด
3	ชะตาชีวิต	หยุด
4	กะทันหัน	หยุด
5	ฝน	หยุด

เริ่มคำสั่งเสียง

ส่งคำสั่งเสียง

กระหึ่มเพลเยอร์

รายการคำสั่ง

ลำดับ	ชื่อเพลง	สถานะ
1	ไกลแค่ไหนคือใกล้	หยุด
2	คู่กัน	หยุด
3	ชะตาชีวิต	หยุด
4	กะทันหัน	หยุด
5	ฝน	หยุด

เริ่มคำสั่งเสียง ส่งคำสั่งเสียง

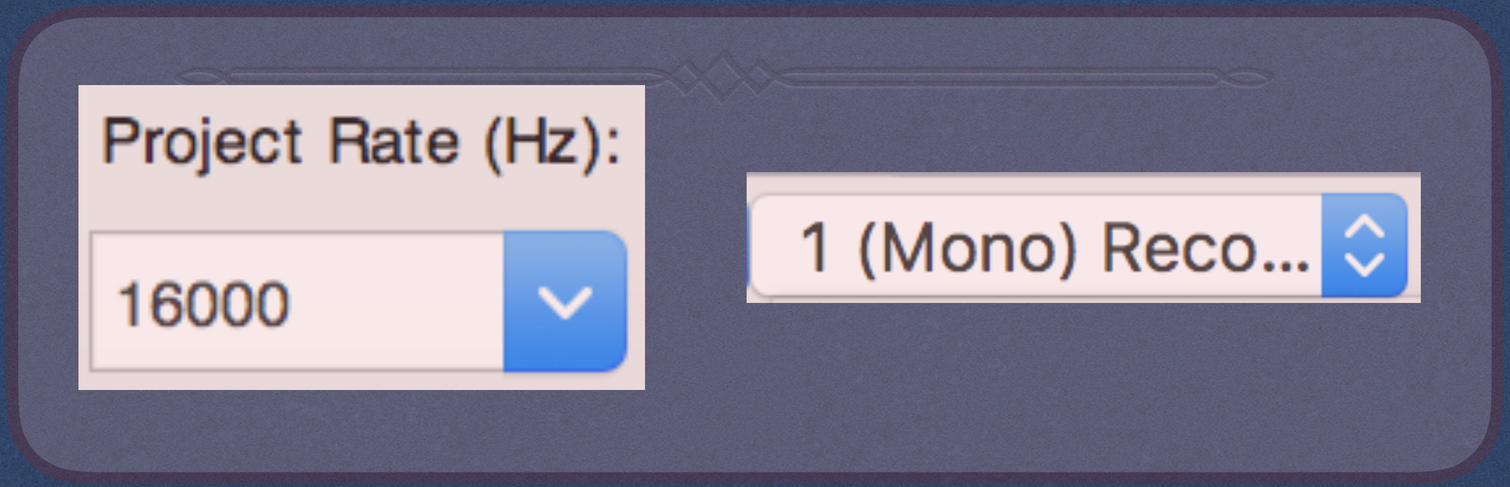
What didn’t worked ?

and mistake#@#$$@

Step 3 - Decoding

ผิด Format !

```
steps/online/nnet2/decode.sh --config conf/decode.config --cmd utils/run.pl \
  --nj 4 --per-utt true --online true exp/nnet2_online/graph \
  data/dev exp/nnet2_online/decode_dev
```



อัดใหม่ -> ลดประโยค -> ประสิทธิภาพลดลง

034	เล่น เพลง ไกลแค่ไหนคือใกล้
035	เล่น เพลง คู่กัน
036	เล่น เพลง ชะตาชีวิต
037	เล่น เพลง กะทันหัน
038	เล่น เพลง ฝน
04	เล่น ไกลแค่ไหนคือใกล้
05	เล่น คู่กัน
06	เล่น ชะตาชีวิต
07	เล่น กะทันหัน
08	เล่น ฝน
1	หยุด
10	หยุด เล่น
1034	หยุด เล่น เพลง ไกลแค่ไหนคือใกล้
1035	หยุด เล่น เพลง คู่กัน
1036	หยุด เล่น เพลง ชะตาชีวิต
1037	หยุด เล่น เพลง กะทันหัน
1038	หยุด เล่น เพลง ฝน
134	หยุด เพลง ไกลแค่ไหนคือใกล้
135	หยุด เพลง คู่กัน
136	หยุด เพลง ชะตาชีวิต
137	หยุด เพลง กะทันหัน
138	หยุด เพลง ฝน
23	เปลี่ยน เพลง

ลืม Sort !

034_1 data/dev/msound_dev/034_1.wav
04_1 data/dev/msound_dev/04_1 .wav
035_1 data/dev/msound_dev/035_1.wav
05_1 data/dev/msound_dev/05_1.wav
036_1 data/dev/msound_dev/036_1.wav
06_1 data/dev/msound_dev/06_1.wav
037_1 data/dev/msound_dev/037_1.wav
07_1 data/dev/msound_dev/07_1.wav
038_1 data/dev/msound_dev/038_1.wav
08_1 data/dev/msound_dev/08_1.wav

dictate.js

A small Javascript library for browser-based real-time speech recognition, which uses Recorderjs for audio capture, and a WebSocket connection to the Kaldi GStreamer server for speech recognition.

```
python -m SimpleHTTPServer
```

```
npm install http-server -g
```

```
http-server -p 8000
```

What worked ?

กระทิ่มเพลเยอร์

รายการคำสั่ง

ลำดับ	ชื่อเพลง	สถานะ
1	ไกลแค่ไหนคือใกล้	หยุด
2	คู่กัน	หยุด
3	ชะตาชีวิต	หยุด
4	กะทันหัน	หยุด
5	ฝน	หยุด

เริ่มคำสั่งเสียง

ส่งคำสั่งเสียง

หยุด เล่น เพลง ฝน

รายการคำสั่ง	หยุด เล่น เพลง ชะตาชีวิต.

หยุด เล่น ไกลแค่ไหนคือใกล้

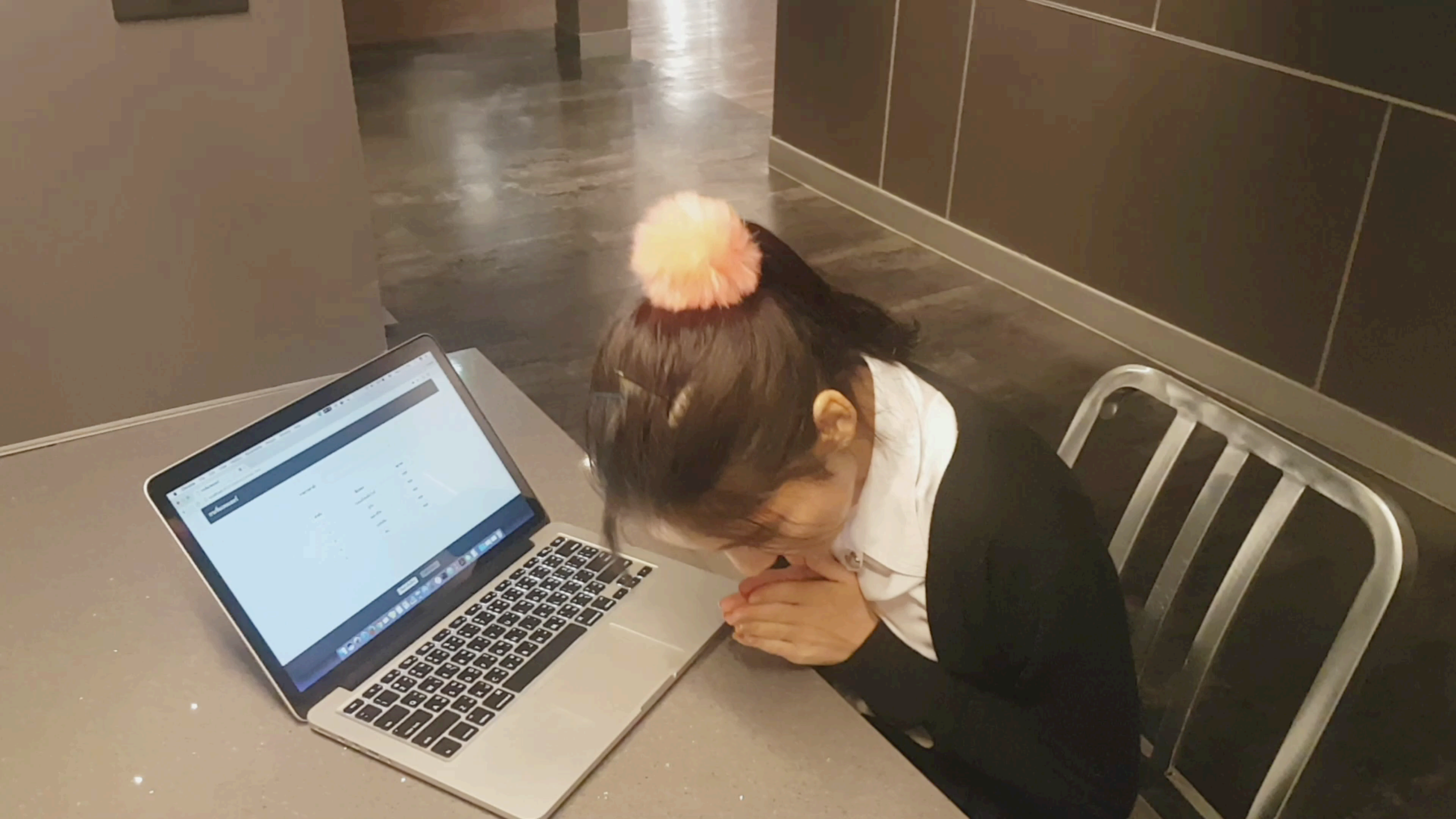
รายการคำสั่ง	เปลี่ยน เล่น ไกลแค่ไหนคือใกล้.

```javascript
function textaction(){
  var order = document.getElementById("trans").value;
  order = order.replace(".", "");
  var mtextSplit = order.split(" ");
  var trainOrder = [["เล่น", "เพลง", "ไกลแค่ไหนคือใกล้"],
                    ["เล่น", "ไกลแค่ไหนคือใกล้"],
                    ["เล่น", "เพลง", "คู่กัน"],
                    ["เล่น", "คู่กัน"],
                    ["เล่น", "เพลง", "ชะตาชีวิต"],
                    ["เล่น", "ชะตาชีวิต"],
                    ["เล่น", "เพลง", "กะทันหัน"],
                    ["เล่น", "กะทันหัน"],
                    ["เล่น", "เพลงฝน"],
                    ["เล่น", "ฝน"],
                    ["หยุด", "เล่น", "เพลง", "ไกลแค่ไหนคือใกล้"],
                    ["หยุด", "เพลง", "ไกลแค่ไหนคือใกล้"],
                    ["หยุด", "เล่น", "เพลง", "คู่กัน"],
                    ["หยุด", "เพลง", "คู่กัน"],
                    ["หยุด", "เล่น", "เพลง", "ชะตาชีวิต"],
                    ["หยุด", "เพลง", "ชะตาชีวิต"],
                    ["หยุด", "เล่น", "เพลง", "กะทันหัน"],
                    ["หยุด", "เพลง", "กะทันหัน"],
                    ["หยุด", "เล่น", "เพลง", "ฝน"],
                    ["หยุด", "เพลง", "ฝน"],
                    ["หยุด", "เล่น"],
                    ["หยุด"],
                    ["เปลี่ยน", "เพลง"]];
```

```javascript
var indexOrder = new Array();
var startOrder = new Array();
var mtext = "";
var startMtext = 10000;
for (var i = 0; i < trainOrder.length; i++) {
  indexOrder.push(0);
  startOrder.push(0);
}

for (var i = 0; i < mtextSplit.length; i++) {
  for (var j = 0; j < trainOrder.length; j++) {
    if (trainOrder[j][indexOrder[j]] == mtextSplit[i]) {
      indexOrder[j]++;
      if (indexOrder[j] == 1)
        startOrder[j] = i;
    }

    if (indexOrder[j] == trainOrder[j].length
        && startOrder[j] <= startMtext) {
      mtext = "";
      startMtext = startOrder[j];
      for (var k = 0; k < trainOrder[j].length; k++) {
        mtext += trainOrder[j][k];
      }
      indexOrder[j]++;
    }
  }
}
```

#$%@#%#$^@#^%@^%#^$#%$^#$%^$

รายการคำสั่ง	คู่กัน. หยุด เล่น กะทันหัน เปลี่ยน หยุด.