Proceedings of the 23rd Conference of the Oriental COCOSDA
November 5-7, 2020, Yangon, Myanmar

Formosa Speech Recognition Challenge 2020 and Taiwanese Across Taiwan Corpus

Yuan-Fu Liao
National Taipei University of Technology
Taipei, Taiwan
yfliao@ntut.edu.tw

Chia-Yu Chang
National Taipei University of Technology
Taipei, Taiwan
chiayu@speech.ntut.edu.tw

Hak-khiam Tiun
National Taitung University
Taitung, Taiwan
hakkhiam@gmail.com

Huang-Lan Su
National Taitung University
Taitung, Taiwan
suhuanglan@gmail.com

Hui-Lu Khoo
National Taiwan Normal University
Taipei, Taiwan
hsuhj@ntnu.edu.tw

Jane S. Tsay
National Chung Cheng University
Chia-Yi Min-Hsiung, Taiwan
Lngtsay@ccu.edu.tw

Le-Kun Tan
National Cheng Kung University
Tainan, Taiwan
lekun1226@gmail.com

Peter Kang
National Donghwa University
Hualien, Taiwan
kang@gms.ndhu.edu.tw

Tsun-guan Thiann
National Taichung University of Education
Taichung, Taiwan
chungoan@mail.ntcu.edu.tw

Un-Gian Iunn
National Taichung University of Education
Taichung, Taiwan
ungian@ntcu.edu.tw

Jyh-Her Yang
Telecom. Lab., Chunghwa Telecom
Taoyuan, Taiwan
houseyang0204@cht.com.tw

Chih-Neng Liang
Automotive Research & Testing Center
Changhua, Taiwan
cnl@artc.org.tw

***Abstract*—Taiwanese (a.k.a. Taiwanese Hokkien, Hoklo, Taigi, Southern Min or Min-Nan) is an endangered language, because the domination of Mandarin, the number of Taiwanese speakers continues to drop, especially among the youth generations. In addressing this problem, a Taiwanese speech-enabled human-computer interface for supporting people's daily life is essential. Therefore, a Formosa Speech in the Wild (FSW) project was established to collect a large-scale Taiwanese speech across Taiwan (TAT) corpus to boost the development of Taiwanese speech recognition (TSR). A Formosa Speech Recognition Challenge 2020 (FSR-2020) was also hosted to promote the corpus as well as to evaluate the performance of state-of-the-art TSR systems. This paper briefly introduces TAT corpus and FSR-2020 challenge, presents the provided data profile, evaluation plan and reports experimental baseline results. A subset of TAT corpus, TAT-Vol1, is given away for free for all participants (non-commercial license), and its corresponding Kaldi baseline recipes have been published online. Experimental results have showed that the combination of TAT corpus and the baseline recipes is a good resource pack for TSR research and development.**

***Index Terms*—Taiwanese across Taiwan Corpus, Formosa Speech Recognition Challenge 2020, Taiwanese Speech Recognition, Machine learning**

I. Introduction

Taiwanese (a.k.a. Taiwanese Hokkien, Hoklo, Taigi, Southern Min or Min-Nan) [1] is in danger of extinction, because the use of Taiwanese was discouraged by the Kuomintang until the 1980s [2], through measures such as assigning Mandarin as the only official language, i.e., ”國語”, banning Taiwanese's use in schools or formal occasions and limiting the hours of Taiwanese broadcast radio or TV programs. The consequence is that Mandarin became dominant in Taiwan and although, nowadays, about 70% of the population still could use Taiwanese [3], most people, especially among the young generations, only have limited vocabulary and cannot speak Taiwanese fluently.

More specifically, a Taiwanese language television channel (Taigi) of the Public Television Service (PTS) [4] was not established until 2019 after the Taiwan National Language Development Act finally gave all of the languages spoken in Taiwan equal status as the country's official national languages in 2018. Moreover, one popular Taigi's programs, the ”全家有智慧” (Smart Family) quiz show, which takes the fun-learning Taiwanese language puzzle competition as the main axis, in fact deeply reflects the miserable situation of Taiwanese language. Because most guests and audiences of the show, even for those elderly, have somehow poor Taiwanese language ability and often have to switch back to Mandarin for smoother communication in the TV show.

It is believed that it is not too late to save this language. To promote Taiwanese language, a Taiwanese speech-enabled human-computer interface, such as Taiwanese voice command, spoken dialogue, speech synthesis or even automatic TV show subtitling or translation systems, for supporting people's daily life is necessary. However, there is no public large-scale

978-1-7281-9896-5/20/$31.00 © 2020 IEEE

Taiwanese speech corpus available for building a state-of-the-art (SOTA) Taiwanese speech recognizer, especially for those deep neural network-based approaches (usually more than 100 hours is required).

In the first phase of FSW [5] project (from 2017 to 2020), a large-scale (3,200 hours) National Education Radio (NER) broadcast Taiwanese Mandarin speech corpus [6], i.e. NER-Trs-Vol1∼17 and NER-Pro-Vol1∼4, had been transcribed and publicly released. By the help of this large NER corpus, a SOTA Mandarin speech recognizer had been built and many Mandarin speech-enabled applications, especially a real-time subtitle generation system for Taiwan Centers for Disease Control (CECC) COVID-19 press conference [7] had been successfully deployed recently.

Being inspired by previous success, in the second phase of FSW project (from 2019 to 2011), we switched our focus to collect a large-scale (more than 300 hours) Taiwanese across Taiwan (TAT) corpus. The results will be used as an infrastructure to boost the research and development of Taiwanese speech recognition (TSR). However, transcription of Taiwanese speech is more difficult than Mandarin due the following issues:

- too many regional variations
- no standard writing system
- only few people can speak Taiwanese fluently
- only few well-trained experts are capable of phonetically transcribing a given Taiwanese speech

Therefore, instead of transcribing spontaneous Taiwanese speech, an alternative strategy was chosen to alleviate those difficulties:

- recruit native Taiwanese speakers across Taiwan to cover regional variations
- adapt the official Taiwanese Romanization system [8] proposed by Taiwan Ministry of Education (MOE) as the writing standard.
- record reading speech with prepared prompt sheets instead of transcribing spontaneous speech by linguists

Currently, the first part (100 hours) of TAT corpus, TAT-Vol1∼2, has been finished. Therefore, we would like to host a FSR-2020 challenge to promote the corpus as well as to evaluate the performance of potential state-of-the-art TSR systems. So, we are now calling for and welcome participants from both academic and industrial sectors to FSR-2020 [9].

In the following sections, TAT corpus and the FSR-2020 challenge will be briefly introduced. Especially, the statistics of TAT-Vol1∼2 will be given and the details of the FSR-2020 including the provided data profile, evaluation plan and the experimental baseline results will be reported. It is believed the combination of TAT corpus and baseline recipe [10] is a good resource pack for further TSR research and development.

II. Background

Taiwanese is a branched-off variety of Southern Min dialects. However, due to history and geographical separation, Taiwanese developed independently from those in Fujian and has many notable differences from those in mainland China. Many of the differences can be attributed to the influence from the languages of Formosan, Dutch and Japanese [1].

A. Regional variations

There are a number of pronunciation and lexical differences between the Taiwanese variants. Some scholars have divided Taiwanese into five subdialects [11] based on geographic region (as shown in Fig. 1) including:

- hái-kháu (海口腔): west coast, formerly referred to as Quanzhou dialect (represented by the Lukang accent)
- phian-hái (偏海腔): coastal (represented by the Nanliao (南寮) accent)
- lāi-po(内埔腔): western inner plain, mountain regions, based on the Zhangzhou dialect (represented by the Yilan accent)
- phian-lāi (偏内腔): interior (represented by the Taibao accent)
- thong-hêng (通行腔): common accents (represented by the Taipei (spec. Datong) accent in the north and the Tainan accent in the south)

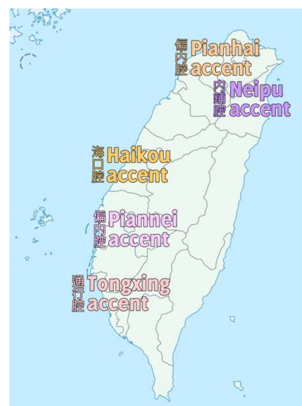


Fig. 1. Distribution of Taiwanese sub-dialects around Taiwan.

The official MOE Taiwanese dictionary [12] further specifies these variations to a resolution of eight regions on Taiwan, in addition to Kinmen and Penghu.

B. Scripts and orthographies

Taiwanese does not have a strong written tradition. Until the late 19th century, Taiwanese speakers wrote mostly in classical Chinese (Hàn-jī, 漢字). However, there are various problems relating to the use of Chinese characters to write Taiwanese, since many morphemes (around 15% of running text) are not definitively associated with a particular character.

Therefore, people began to use various different methods to solve this issue. Such as using either a Latin-based script or through the use of a Chinese character chosen phonetically with no relation to the original word via meaning.

Among many systems of writing Taiwanese using Latin characters, the most used system is called peh-oē-jī (POJ, 白話字) developed by Western missionaries in the 19th century. POJ can also be used along with Chinese characters in a mixed script called Hàn-lô (漢羅), where words specific to Taiwanese are written in POJ, and words with associated characters written in Han Characters.

Recently, MOE officially promoted the Taiwanese Romanization System (Tâi-lô, 台羅) to standardize the writing system in 2006. Furthermore, in 2009, MOE formulated and released a list of Taiwanese Southern Min Recommended Characters, in total 700 Han characters, for the use of writing uniquely Taiwanese words in the Hàn-jī or mixed Hàn-lô (漢羅) approaches.

III. Data Collection

The aim is to construct a proper speech corpus for building a high-performance automatic TSR system. To avoid the labor-intensive Taiwanese speech transcription efforts, in the phase II of FSW project, we chose to use prompt sheets to collect reading speech from 600 Taiwanese native speakers around Taiwan and maximize the divergence in text, gender, age, speaker and environment condition under the constrain of financial funding. Here, the text articles for compiling prompt sheets written in native POJ, Tâi-lô or Hàn-lô are preferred than Chinese orthography for precise Taiwanese speech recording and different microphone configurations were utilized to cope with all types of audio qualities to simulate a day-to-day audio recording condition.

A. Corpus Design

1) Text Materials and Prompt Sheets: The adopt Taiwanese native text articles mainly came from articles published by two sources:

- Li Kang Khioh Taiwanese Cultural and Educational Foundation (李江台語文教基金會) [13]: 50 authors, about 6,000 words per author, and a daily conversation textbook with 14 lessons
- MOE: 250 articles, about 600 words per article

Use of these articles has been authorized by the authors.

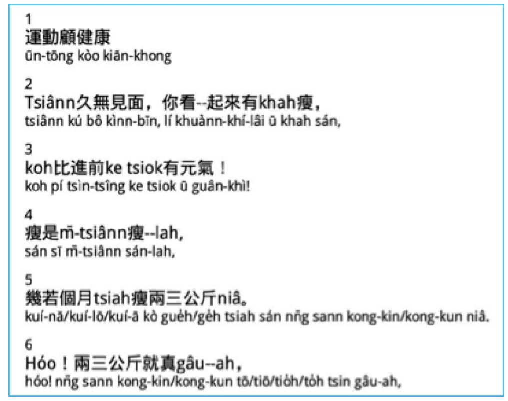
To cover all possible language usages, three sessions were designed in the prompt sheets including:

- Numbers (date, address, phone, ID, $\cdots$)
- Daily conversations
- Short articles

All materials are presented sentence-by-sentence (or short phrase-by-phrase). Moreover, Tâi-lô scripts are also given for references. A typical ”Daily Conversation” session (partial) is shown in 2.

2) Speaker Distribution: The first part of TAT corpus was recorded from 200 native speakers of Taiwanese. Each speaker produced about 30 minutes speech. To reflect the current status of Taiwanese speech, these speakers were recruited across Taiwan and covered different ages, especially kids, youths and elderly.

To this end, 5 collaborators based in Taipei (in Northern Taiwan), Taichung (two teams, in central Taiwan), Chiayi and Tainan (in southern Taiwan) were asked to recruit Taiwanese native speakers from their neighbourhood (about 40 speakers per site).


1
運動顧健康
ūn-tōng kòo kiān-khong

2
Tsiânn久無見面，你看--起來有khah瘦，
tsiânn kú bô kìnn-bīn, lí khuànn-khí-lâi ū khah sán,

3
koh比進前ke tsiok有元氣！
koh pí tsìn-tsîng ke tsiok ū guân-khì!

4
瘦是m̄-tsiânn瘦--lah，
sán sī m̄-tsiânn sán-lah,

5
幾若個月tsiah瘦兩三公斤niâ。
kuí-nā/kuí-lō/kuí-ā kò gueh/geh tsiah sán nn̄g sann kong-kin/kong-kun niâ.

6
Hóo！兩三公斤就真gâu--ah，
hóo! nn̄g sann kong-kin/kong-kun tō/tiō/tioh/toh tsin gâu-ah,

Fig. 2. The daily conversation session of a typical prompt sheet used during recording.

B. Recording Protocol

1) Microphone Configurations: Six different microphones were adopted at the same time for data collection in a quiet office environment as shown in Fig. 3.

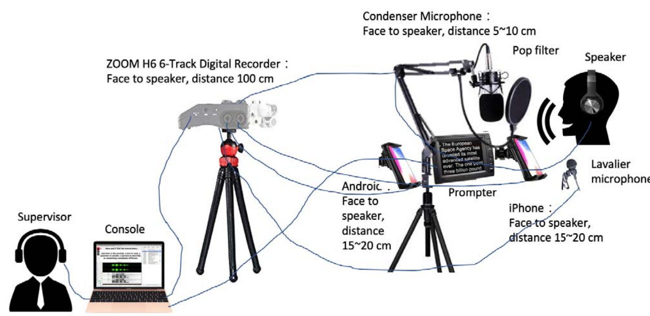


Fig. 3. Schematic diagram of the configuration of the recording equipment.

Specifically, we have three microphones—one close-talk, one collar clip-on and one distant X-Y stereo microphone to pick up the direct, indirect near-field mono sounds and far-field stereo image, respectively. Additionally, we have one iPhone and one Android phone that acted as live microphone to simulate a more day-to-day audio recording condition.

These six microphones were connected to a multi-track digital audio interface for live recording. The audios, in total six channels, were then synchronized and recorded in a six-track WAV format, with a sampling rate of 96 kHz and a bit resolution of 24 bits. Here are the details of the digital audio interface and microphones used for recording:

- Digital audio interface: ZOOM H6
- Close-talk: Audio-Technica AT2020
- Lavalier: Superlux WO518+PS418D
- Distant: ZOOM XYH-6 stereo microphone
- iPhone App: Microphone Live
- Android Phone App: Microphone

2) Recording Software: SpeechRecorder [14] was customized to display Taiwanese text and Tâi-lô scripts and to capture six audio tracks at the same time. The main feature of

SpeechRecorder is that the speaker and supervisor will have different screens and views. Therefore, the speaker will not be distracted during recording.

IV. Statistics of TAT Corpus

The recorded 6-track speech data, were first separated into 6 single channel signals and then converted into WAV format with 16 kHz sampling rate with 16-bit PCM encoding. Some statistics of TAT corpus are reported in the following subsections.

A. Numbers of Speakers, Sentences and Characters

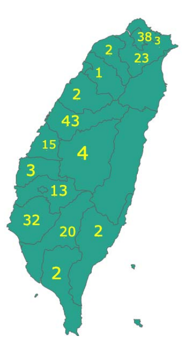
For training a Taiwanese speech recognizer, the collected speech data was further splitted into 2 volumes, i.e., TAT-Vol1∼2 and divided into the train, evaluation and test subsets by 8:1:1. Fig. 4 shows the content of TAT-Vol1∼2 including number of speakers, sentences, characters, and total speech duration in hours. TAT-Vol1∼2 had been published online using a GitLab server for easier issue report and error correction.

TAT-Vol1				
	Speakers	Sentences	Characters	Hours
Train	80	23,104	271,772	41.76
Evaluation	10	2,943	34,426	5.02
Test	10	2,786	33,394	5.16
TAT-Vol2				
	Speakers	Sentences	Characters	Hours
Train	80	23,216	272,671	42.39
Evaluation	10	2,951	35,951	4.76
Test	10	2,811	31,985	5.27
Total				
	Speakers	Sentences	Characters	Hours
Total	200	57,811	680,199	104.36

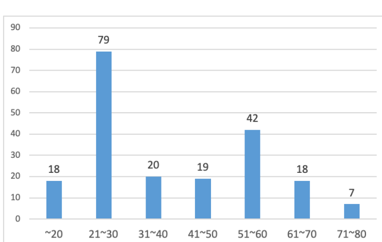
Fig. 4. Numbers of Speakers, Sentences and Characters in TAT corpus.

B. Distribution of Speakers' Genders, Residences and Ages

There are in total 90 male and 110 female speakers in TAT-Vol1∼2 corpus. Fig. 5 shows the distributions of speakers' current residences and ages. From the figure, it was found that we need to recruited more elderly speakers and speakers lived in the east Taiwan. This issue will be fixed in the second and third year (2020∼2021) of FSW project.



(a)



(b)

Fig. 5. Distributions of the (a) current residences and (b) ages of the recruited 200 speakers in TAT corpus.

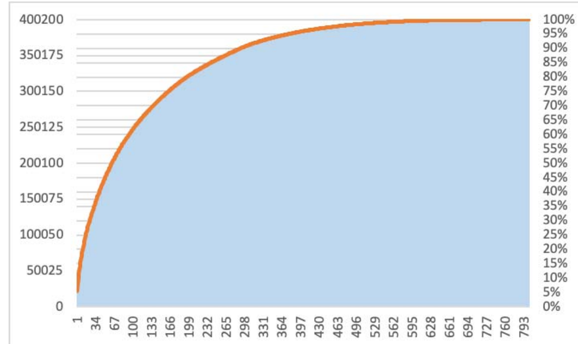
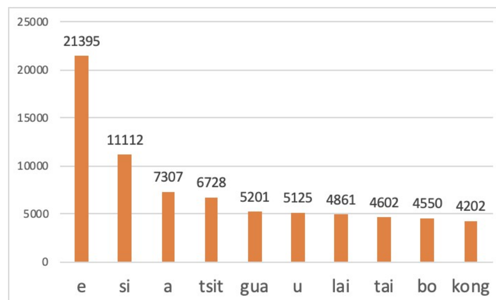


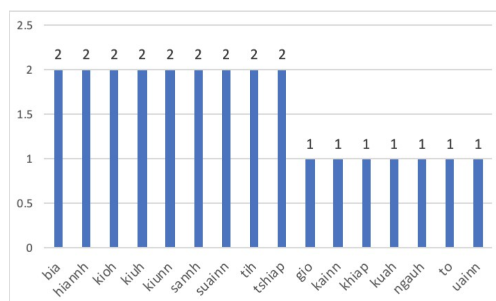
Fig. 6. Accumulated histogram of the syllables in TAT corpus.

C. Accumulated Histogram of Syllables

Fig. 6 shows the accumulated Taiwanese syllables frequency in TAT-Vol1∼2 corpus. On the other hand, Fig. 7 shows the highest- and lowest-frequency (count<3) syllables. There are in total 803 legitimate Taiwanese syllables. From those figures, it was found that (1) 750 out of 803 syllables had been covered, (2) 388 syllables could cover 95% of the running speech and (3) /e/ and /si/ are the highest-frequency syllables. The issue of 53 missing syllables should be fixed as soon as possible.



(a)



(b)

Fig. 7. The (a) highest- and (b) lowest-frequency syllables in TAT corpus.

V. Formosa Speech Recognition Challenge 2020

A. Task

The main task of FSR-2020 is to built a Taiwanese speech recognizer that could output either:

1) Traditional Chinese characters (繁體中文字), i.e., Taiwanese speech to Chinese Characters

TABLE I
THE TAIWANESE ROMANIZATION LEXICON (PARTIAL) PROVIDED BY FSR-2020 COMPETITION.

Index	Syllable	Initial	Final
1	an	iNull	an
2	bah	b	ah
3	jiok	j	iok
4	ngiau	ng	iau
...	...	...	...

2) Taiwanese Southern Min Recommended Characters (漢字) proposed by Taiwan MOE
3) Tâi-lô (台羅) scripts in Taiwanese Romanization

Here shows some typical examples of the required system outputs:

1) 現在是晚上八點
2) 這馬是暗時八點
3) tsit4 ma2 si7 am3 si5 peh4 tiam2

FSW-2020 is featured with (1) a free Taiwanese speech corpus, i.e., TAT-Vol1, and a Tâi-lô syllable lexicon, (2) a reference baseline recipe, and (3) two-stage (pilot-test and final-test) evaluation. Since the purpose of this challenge is to boost the development of Taiwanese-specific techniques, the pilot-test is set up as a warm-up exercise only.

B. Evaluation metric

The performance will be evaluated using character or syllable error rate (CER or SER in %). Since some characters pairs are exchangeable, for example,"台北", "臺北". In these cases, either will be fine. A script will be written to convert those characters before evaluation.

C. Schedule

FSW-2020 was scheduled as follows:

- 2020/06/01 — Registration Open
- 2020/09/01 — Pilot-Test (dry-run)
- 2020/12/01 — Registration Close
- 2021/01/01 — Final-Test

A more detail plan of the challenge is listed on the challenge website [9].

D. Provided Corpus and Lexicon

FSW-2020 is based on TAT-Vol1 corpus (free for all participants). Moreover, a Taiwanese Romanization lexicon (Tâi-lô syllable to /initial/ and /final/) was also provided. The lexicon has 803 Taiwanese syllables. A typical example of the lexicon are shown in Table I.

E. Reference Baseline

To demonstrate the development of a TSR system, a reference baseline was and released as a reference. The baseline system, alias the "Taiwanese Speech Recognition Recipe", was built using the Kaldi toolkit [15] and a set of Unix shell scripts. This recipe closely followed our previous kaldi "formosa" example scripts. It will train 6 Hidden Markov Model/Gaussian Mixture Model (HMM/GMM)- and 2 hybrid HMM/time delay neural network (HMM/TDNN) systems. Among them, only the results of the best system, i.e., the Chain/TDNN model, was treated as the baseline system and reported here.

TABLE II
THE RECOGNITION PERFORMANCE (SER IN %) OF THE BASELINE SYSTEM EVALUATED ON TAT-VOL1-EVAL AND TAT-VOL1-TEST SUBSETS.

Database	TAT-Vol1-eval	TAT-Vol1-test
SER	12.81	11.24

Fig. 8 shows the architecture of the provided Chain/TDNN-based acoustic model. Its input feature vector included a 229-dimensional feature vector that was composed of three 43-dimensional MFCCs and pitch features plus a 100-dimensional i-vector. It had 6 hidden layers and 626 neurons per-layer. From the low- to top-level layers, the splicing windows of the hidden layers are set to (–1,0,1), (–1,0,1), (–3,0,3), (–3,0,3) and (–3,0,3). Speed perturbation-based data augmentation approach was also applied to increase the number of training samples by 3 times.

Finally, during decoding, a syllable-level tri-gram language model trained with the transcriptions of TAT-Vol1-train corpus was utilized. Table II shows the experimental results of the baseline system trained on TAT-Vol1-train and evaluated on TAT-Vol1-eval and TAT-Vol1-test corpus. From the table, it is found that about 11.24% SER was achieved. This reveals the effectiveness of the provided baseline system.

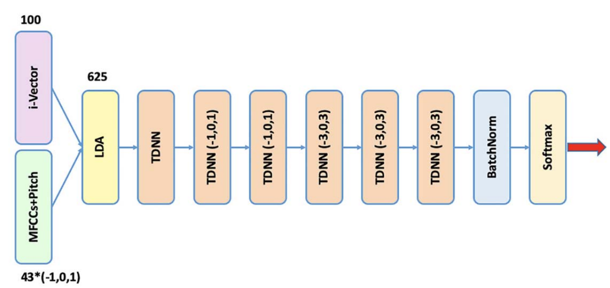


Fig. 8. The architecture of the deep time delay neural network (TDNN)-based acoustic model adopt in the FSR-2020 reference baseline system.

VI. ADVANCED TSR

Here further shows the performance of the hybrid HMM/TDNN-based system retrained using more data (called "advanced TSR" from now on) and evaluated on difference microphone or speaking style (reading vs. spontaneous speech).

1) Different Microphones: To understand the influence of different microphones, two different experimental settings were applied here including:

- Single microphone: training using data from only a single microphone
- All six microphone: training size increases by 6 times by using all data from all six microphones

Fig. 9 shows that (1) the performance of near-field condenser and lavalier microphones were the best and the distant XYH-

TABLE III
THE RECOGNITION PERFORMANCE (SER IN %) OF THE ADVANCED TSR SYSTEM EVALUATED ON PTS CORPUS.

Case	1	2	3
Database	TAT-Vol1∼2	PTS	TAT-Vol1∼2+PTS
Hours	480	87	567
SER	44.63	28.91	21.50

X and XYH-Y were the worst and (2) advanced TSR system trained using speech from multiple microphones did increase the robustness of the TSR system, since the average SER greatly reduced from 7.15% to 5.02%. These results confirmed the necessity of multiple microphone recordings.

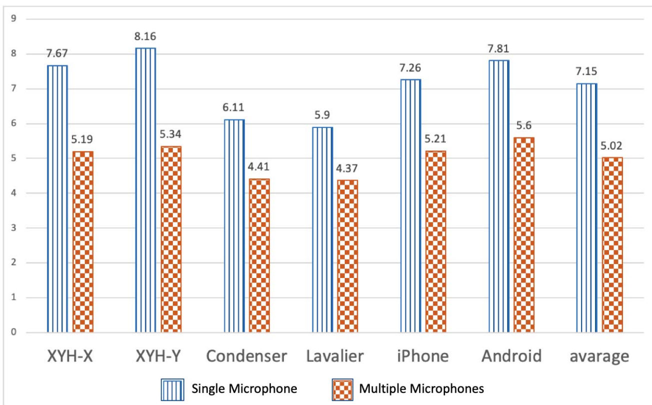


Fig. 9. Performance of advanced TSR system on different microphones on TAT corpus.

2) Different Speaking Style: On the other hand, the advanced TSR system was evaluated on an extra proprietary spontaneous speech database collected from PTS' broadcast news and drama shows (called PTS corpus from now on). There are in total about 87 and 27 hours speech data in the PTS training and test subsets, respectively. Three different configurations were tested as follows:

1) trained on TAT-Vol1∼2, tested on PTS
2) trained on PTS-train, tested on PTS
3) trained on TAT-Vol1∼2+PTS, tested on PTS

Table III shows all the experimental results. First, it was found that, in case 1, the performance of the advanced TSR system is only 44.63%. Since this is a worst test case (mismatch speaking style) and only syllable-level tri-gram was applied, we may consider that this performance is still acceptable. Secondary, in case 2, better SER (28.91%) could be achieved by retraining the advanced TSR system using PTS data. This indicated the importance of training and testing on match data. Finally, if both corpora were combined to retrain the advanced TSR system (case 3), its performance could be greatly improved from SER 28.91% to 21.59%. This results indicated that our prompt sheet-based data collection strategy is helpful, but should somehow adjust it a little bit to record more spontaneous speech in the future for daily-life speech-enable applications.

VII. CONCLUSIONS

This paper introduces the construction of a large-scale TAT corpus and an on-going FSR-2020 challenge in response to an arising interest in TSR technologies. After one year of hard work, we had collected, in total, about 100 hours Taiwanese reading speech using six different microphones from 200 speakers across Taiwan. The first volume of TAT corpus, i.e., TAT-Vol1, and a corresponding Kaldi script-based reference TSR baseline were released to support the FSR-2020 competitions. So far, there are already more than 20 registered challenge participants. It is hope that TAT and FSR-2020 will become a solid infrastructure for boosting the development of TSR techniques.

ACKNOWLEDGEMENTS

This work is supported partially by Taiwan's Ministry of Education under project "教育部閩南語語音語料庫建置計劃" and partially by Ministry of Science and Technology under contract No. 107-2221-E-027-102, 107-2911-I-027-501, 107-3011-F-027-003, 108-2221-E-027-067, 109-2221-E-027-108, partially by Telecom. Lab., Chunghwa Telecom, Taoyuan Taiwan under contract No. TL-108-D303. and partially by the Department of Industrial Technology of Ministry of Economic Affairs under contract No. 109-EC-17-D-11-1679.

REFERENCES

[1] "Taiwanese Hokkien - Wikipedia." [Online]. Available: https://en.wikipedia.org/wiki/Taiwanese_Hokkien
[2] "Martial law in Taiwan - Wikipedia." [Online]. Available: https://en.wikipedia.org/wiki/Martial_law_in_Taiwan
[3] Taiwan Accounting and Statistics, "Languages used at home for the resident nationals - 2010 Population and Housing Census," Tech. Rep., 2010. [Online]. Available: https://ebas1.ebas.gov.tw/phc2010/english/51/a1.pdf
[4] "PTS Taiwan." [Online]. Available: http://eng.pts.org.tw/
[5] Yuan-Fu Liao, "Formosa Speech in the Wild Project." [Online]. Available: https://sites.google.com/speech.ntut.edu.tw/fsw
[6] Y. F. Liao, Y. H. S. Chang, Y. C. Lin, W. H. Hsu, M. Pleva, and J. Juhar, "Formosa Speech in the Wild Corpus for Improving Taiwanese Mandarin Speech-Enabled Human-Computer Interaction," *Journal of Signal Processing Systems*, aug 2019.
[7] Ministry of Science and Technology Center for Global Affairs and Science Engagement, "Speech Recognition Technology Contributed to Taiwan's Successful Pandemic Control." [Online]. Available: http://gase.most.ntu.edu.tw/taiwancanhelp/national_item/24#!
[8] Ministry of Education, Taiwan, "臺灣閩南語羅馬字拼音方案使用手冊," Tech. Rep., 2008.
[9] Yuan-Fu Liao, "Formosa Speech Recognition Challenge 2020 - Taiwanese ASR." [Online]. Available: https://sites.google.com/speech.ntut.edu.tw/fsw/home/challenge-2020
[10] Chia-Yu Chang, "GitHub - FSR-2020 Baseline." [Online]. Available: https://github.com/t108368084/NTUT-TAT-Baseline
[11] H. Klöter, "Táiwān: Language Situation," *Encyclopedia of Chinese Language and Linguistics*, vol. 4, pp. 263–267, 2016. [Online]. Available: https://www.academia.edu/38138351/Taiwan_Language_Situation
[12] Ministry of Education, Taiwan, "臺灣閩南語常用詞辭典." [Online]. Available: https://twblg.dict.edu.tw/holodict_new/index/fulu_fangyan_level1.jsp
[13] "Li kang khioh taiwanese cultural and educational foundation (李江台語文教基金會)." [Online]. Available: https://www.tgb.org.tw/
[14] "SpeechRecorder." [Online]. Available: https://www.bas.uni-muenchen.de/Bas/software/speechrecorder/
[15] G. kaldi asr/kaldi, "Kaldi Speech Recognition Toolkit," 2016. [Online]. Available: https://github.com/kaldi-asr/kaldi