Proceedings of the 12th Conference on Language Resources and Evaluation (LREC 2020), pages 6593–6598
Marseille, 11–16 May 2020

On Construction of the ASR-oriented Indian English Pronunciation Dictionary

Xian Huang, Xin Jin, Qike Li, Keliang Zhang
Information Engineering University
Luoyang City, Henan Province, China
a77huang852yi@126.com, 630530474@qq.com, 1037038328@qq.com, kliang99@sina.com

Abstract

As a World English, a New English and a regional variety of English, Indian English (IE) has developed its own distinctive characteristics, especially phonologically, from other varieties of English. An Automatic Speech Recognition (ASR) system simply trained on British English (BE) /American English (AE) speech data and using the BE/AE pronunciation dictionary performs much worse when applied to IE. An applicable IEASR system needs spontaneous IE speech as training materials and a comprehensive, linguistically-guided IE pronunciation dictionary (IEPD) so as to achieve the effective mapping between the acoustic model and language model. This research builds a small IE spontaneous speech corpus, analyzes and summarizes the phonological variation features of IE, comes up with an IE phoneme set and complies the IEPD (including a common-English-word list, an Indian-word list, an acronym list and an affix list). Finally, two ASR systems are trained with 120 hours IE spontaneous speech data, using the IEPD we construct in this study and CMUdict separately. The two systems are tested with 50 audio clips of IE spontaneous speech. The result shows the system trained with IEPD performs better than the one trained with CMUdict with WER being 15.63% lower on the test data.

Keywords: Indian English, phonological variation features, pronunciation dictionary

1. Introduction

Indian English (IE) refers to the English language used in the Republic of India, a South Asian nation and member of the commonwealth (McArthur, 2001:290). IE is a "World English" (Kachru, 1985), a "New English" (Platt et al., 1984:2-3), used in the outer circle of the three concentric circles of English (Kachru, 1985:12-15), and is a regional variety of English. In India, English is widely used as a second language — it is the associate language of India, the language of science and technology, one in the "three language formula" of the Indian education system, and also the language of administration and the armed forces, and is used as the link language among people who do not share a common Indian language.

The English used in India has developed its own linguistic patterns as a result of the influences of indigenous Indian languages and the socio-cultural background in the country. The deviation from native English are even greater with regard to phonological and phonetic patterns, though differences in lexis and grammar also exist. Table 1 is summarized by Kachru (1983:84) of Bansal (1969) on the intelligibility of IE.

Participants	Highest (%)	Lowest (%)	Average (%)
1 IE & Received Pronunciation (RP) speakers (group)	73	67	70
2 IE & RP speakers (cline of intelligibility)	95	53	
3 IE &AE speakers	81	72	74
4 IE &German speakers	67	40	57
5 IE speakers & Nigerians	66	34	53
6 IE speakers with other IE speakers	88	54	74
7 RP speakers with other RP speakers	100	95	97

Table 1: Intelligibility of IE

From Table 1 we can see the intelligibility of Indians forms a cline, with those of high English proficiency speaking almost like the natives, while the other end of the cline causing great troubles to the listeners for comprehension. As Gramley & Patzold (1992:441) mentioned, "there is a great deal of local variation which depends especially on the native language of any given IndE user and is further influenced by spelling pronunciations; even within India mutual comprehension cannot always be expected", not to mention people speaking other varieties of English.

Compared with the ASR of American English (AE) or British English (BE), few studies are done as to the speech recognition of IE. Amodei, et al(2016) used 11940 hours AE speech data as the training set and built an ASR system without using pronunciation dictionaries. Then the system is used to transcribe different English varieties. The result shows that when dealing with American read material, the word error rate (WER) was 3.1% and when dealing with American/Canadian spontaneous speech, WER was 7.94%. However, when the system was used to transcribe IE, WER was much higher and was 22.89%. Pull & Kumar (2016) also finds out that it is necessary to use IE speech data to train the acoustic model.

From above introduction we can see that an AE/BE based ASR system will definitely be influenced by the local characteristics of the Indian accents. Therefore to build an effective IEASR, IE spontaneous speech data and a pronunciation dictionary which can fully embody the pronunciation characteristics of IE are needed.

Fig. 1 shows the process of construction of a mainstream ASR system (Rabiner, 2011:953):

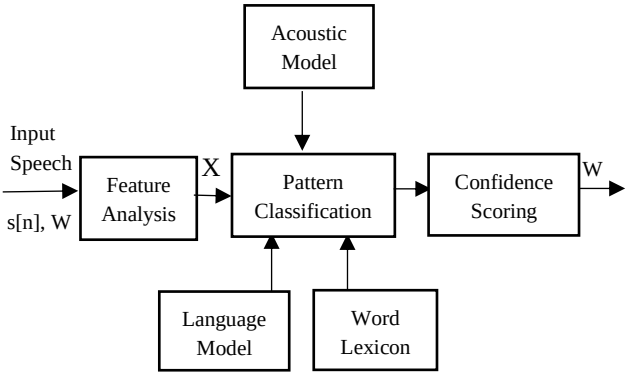

Figure 1: Basic Diagram of Overall Speech Recognition System

From Figure 1 we can see the pronunciation dictionary (word lexicon) contains the mapping between the words and the phones, which is to say, it builds the mapping relationship between units of the acoustic model and the language model, thus connecting the two and forming a searching space for decoding of the decoder.

When it comes to English pronunciation dictionaries, Everyman's English Pronunciation Dictionary (EEPD) edited by Daniel Jones (Version 1-13) and A.C. Gimson (Version 14) is a publicly acknowledged dictionary that best reflects the British Received Pronunciation (BRP). Longman Pronunciation Dictionary (LPD) compiled by J.C.Wells is also a comprehensive and authoritative dictionary of the English pronunciation. The two English pronunciation dictionaries are compiled by authoritative phoneticians, cover a wide range of English words, and include varied forms of words as well as the marks of accents and syllables, thus can offer a very good resource and guidance for the compilation of electronic pronunciation dictionaries.

Large-scale English pronunciation dictionaries oriented for speech technologies and applications are PRONLEX, CELEX, CMUdict and so on. PRONLEX is developed mainly for speech recognition. PRONLEX contains 90,988 lexical entries and includes coverage of WSJ30, WSJ64, Switchboard and CALLHOME English. WSJ30K and WSJ64K are word lists selected from several years of Wall Street Journal texts used in recent ARPA Continuous Speech Recognition corpora. Switchboard is a three-million-word corpus of telephone communications on a variety of topics. The Carnegie Mellon University Pronouncing Dictionary (CMUdict) is an open-source machine-readable pronunciation dictionary for North American English (NAE) that contains over 134,000 word forms and their pronunciations. Its entries are particularly useful for speech recognition and synthesis, as it has mapping from words to their pronunciations in the ARPAbet phoneme set, a standard for English pronunciation. The CELEX pronunciation dictionary enlists 41,000 original-form words of the Oxford Advanced Learner's Dictionary (the 1974 version) and 53,000 original-form words of the Longman Dictionary of Contemporary English (the 1978 version), including in total the pronunciation of 160, 595 word forms. PRONLEX and CMUdict are based on AE pronunciation, annotated with the ARPAbet system; CELEX is based on BRP, its phoneme set being deduced from IPA. PRONLEX, CMUdict and CELEX all divide the stress system into three levels: primary stress, secondary stress and no stress.

To date the only publicly available list of IE pronunciation is offered by Nihalani et al. (1979). The book, based on the Educated Indian English (EIE), offers the Indian Recommended Pronunciation (IRP) of about 2,800 words. Each entry includes the form of the word, the BRP pronunciation as well as the IRP pronunciation, annotated in IPA. However, the number of entries is far too small to satisfy the need of ASR training.

Though the ASR-oriented pronouncing dictionaries mentioned above enlisted large amounts of entries, they cannot be used to build an effective IEASR system, the reasons are as follows:

(1) The dictionaries are either based on BE (CELEX) or AE (CMUdict, PRONLEX) pronunciation. The fact is that the pronunciation of a lot of words underwent changes in IE.

(2) ARPAbet—the annotation system of CMUdict or PROLEX—was designed exclusively for the annotation of AE. Some phonemes (such as /t̪/、/d̪/) in IE are absent in ARPAbet. Thus the pronunciation of some IE words cannot be correctly transcribed.

(3) In Indian people's daily communication there exist several Indian words, such as people's names (Indra, Jamal), names of places (Vellore, Vishakapatname), names of food (parathas), and other proper names (Rajya Sabha). They appear frequently in the speech of Indian people but are absent in the existent English pronouncing dictionaries.

(4) Indian people like to use acronyms, such as BJP (Bharatiya Janata Party), AP (Andhra Pradesh), UP (Utar Pradesh) and so on, which are absent in the above-mentioned dictionaries.

From above discussion we can see there is no comprehensive IE pronunicaiton dictionary which can fully embody the pronunciation of IE till now. This research aims at building a pronunciation dictionary that accords with the actual pronunciation habits of the Indian people. The structure of this study mainly consists of four parts: (1) construction of an IE spontaneous English Speech Corpus; (2) the analysis and summarization of IE phonological variation features; (3) construction of an IEPD; (4) an experiment testing the effect of the IEPD.

2. Construction of an IE Spontaneous Speech Corpus

To fully study the phonology of IE spontaneous speech and compare the differences between the pronunication of NAE and IE, the thesis selects 10 speakers who were born, raised up and received education in New Delhi; what's more, their mother tongue is Hindi and they have no long-time overseas experiences.

The reasons to choose the above-mentioned speakers are as follows:

(1) New Delhi is described as "the heartland of proficiency in English". People who are born and receive education in New Delhi can enjoy ample opportunities for learning English. If their speech also contain accents, then these accents quite probably exist in most Indian people's speech.

(2) Hindi is the official language and the most-widely used indigenous language in India. People whose mother language being Hindi represent a majority of the Indian population.

(3) With no long-time overseas experience, the representatives can speak IE naturally without imitating the native speakers or their accents being influenced or corrected by the native English speakers.

After selecting the representative speakers, we collect their speech in interviews. The interviews are mainly from programs like "The Big Fight" , "NDTV Dialogue" and so on from the NDTV website, and NDTV is a famous and influential Indian English TV station in New Delhi. There are in all 3345 audio clips, amounting to a total length of more than 5 hours. Each audio clip is annotated with Praat, a widely-used software in phonetic analysis. The pronunciation of speech is annotated with Speech Assessment Methods Phonetic Alphabet (SAMPA), a computer-readable phonetic script using 7-bit printable ASCII characters, based on the International Phonetic Alphabet. Transcribing the pronunciation of the speech materials in SAMPA facilitates the retrieval and

comparison of sounds. The speech materials are annotated in seven layers : Transcription (transcribed texts of the speech), Phonetization (Phones of the words in the speech), TokensAlign (Time-Aligned Tokens), PhonTokensAlign-NAE (North American English pronunciation of the aligned tokens), PhonTokensAlign-IE (The actual pronunciation of the aligned tokens), NoSyllable (Actual number of syllable of the token), Comments (phonetic or phonological phenomena that need to pay attention to), containing rich phonetic as well as phonological information. Fig.2 shows the annotation of speech materials in the corpus.

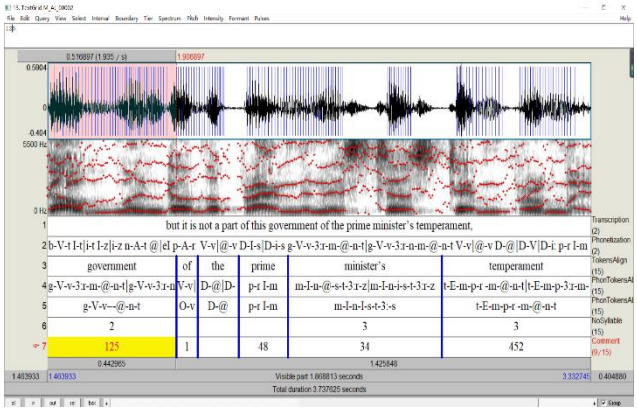

Figure 2: Annotation of the IE Speech Material

Then the annotation is extracted through Python programs and saved in an Excel which facilitates fast concordance, as is shown in Fig.3:

	A	B	C	D	E	F	G	H	I
11132	TokensAlign	PhonTokensAlign-AE	PhonTokensAlign-IE	Former	After	NoSyllable	Comment	Filename	
11133	well	w-E-l	w-E-l					M_AI_00001.TextGrid	
11134	the	D-@|D-V|D-i:	d_d-i:		i-l-E-k_h-		1	M_AI_00001.TextGrid	
11135	election	i-l-E-k-S-@-n	i-l-E-k_h-S-@-n	d_d-i:	k_h-a-l-@	3	4	M_AI_00001.TextGrid	
11136	calendar	k-{-l-@-n-d-3:r	k_h-a-l-@-n-d-3:	i-l-E-k_h-l-s		3	14-	M_AI_00001.TextGrid	
11137	is	I-z|i-z	I-s	k_h-a-l-@-h-E-d				M_AI_00001.TextGrid	
11138	ahead	@-h-E-d	@-h-E-d	I-s	O-f		$	M_AI_00001.TextGrid	
11139	of	V-v|@-v	O-f	@-h-E-d	V-s		3	M_AI_00001.TextGrid	
11140	us	V-s	V-s	O-f	a-d-a-m-			M_AI_00001.TextGrid	
11141	but	b-V-t	b-V-t_h	a-m	I-t_h			M_AI_00002.TextGrid	
11142	it	I-t|i-t	I-t_h	b-V-t_h	I-s			M_AI_00002.TextGrid	
11143	is	I-z|i-z	I-s	I-t_h	n-O-t_h		3	M_AI_00002.TextGrid	
11144	not	n-A-t	n-O-t_h	I-s	E		3	M_AI_00002.TextGrid	
11145	a	@	E	n-O-t_h	p-A-r\-t		1	M_AI_00002.TextGrid	
11146	part	p-A-r\-t	p-A-r\-t	E	O-v		1+$	M_AI_00002.TextGrid	
11147	of	V-v|@-v	O-v	p-A-r\-t	D-I-s		3	M_AI_00002.TextGrid	
11148	this	D-I-s|D-i-s	D-I-s	O-v	g-V-v-m-			M_AI_00002.TextGrid	
11149	government	g-V-v-3:r-m-@-n-t|g-V-v-3r-n-m-@	g-V-v-m-@-n-t	D-I-s	O-v	2	562	M_AI_00002.TextGrid	
11150	of	V-v|@-v	O-v	g-V-v-m-	D-@			M_AI_00002.TextGrid	
11151	the	D-@|D-V|D-i:	D-@	O-v	p-r\-al-m			M_AI_00002.TextGrid	
11152	prime	p-r\-al-m	p-r\-al-m	D-@	m-I-n-I-s		48	M_AI_00002.TextGrid	
11153	minister's	m-I-n-@-s-t-3:r-z|m-I-n-i-s-t-3:r-z	m-I-n-I-s-t-3-s	p-r\-al-m	t-E-m-p-	3	34-	M_AI_00002.TextGrid	
11154	temperament	t-E-m-p-r\-@-m-@-n-t|t-E-m-p-3:r-m	t-E-m-p_h-r\-u-m-@-n-t	m-I-n-I-s		3	452	M_AI_00002.TextGrid	

Figure 3: Annotation of the Speech in Excel

3. Analysis of the Phonological Variation Features of IE

Scholars especially Indian scholars analyze the phonology of IE from different perspectives, with both general IE and regional varieties of IE as research objects. Vowels/consonants, word stress and intonation of IE have been studied in depth. However, facing IEASR tasks, there still exist several problems to be solved: 1) most studies take read speech as research objects while IEASR needs to deal with continuous spontaneous speech; 2) existing studies mainly study the vowels/consonants, word stress and intonation of IE, while phonotactics, phonological alteration and rhythm are seldom studied. However, the latter part is also an important part of phonology and are necessary to be dealt with in ASR tasks.

In this study, we compare the actual pronunication of IE words with the entries in CMUdict, trying to find out the differences between IE and NAE pronunciation. The following are the main findings:

As to the consonants, it is found that in IE word-initial /p, t, k/ are not aspirated; labio-dental fricative /v/ is often pronounced as [ʋ]; dental fricatives /θ, ð/ are often pronounced as [t̪ʰ] and [d̪]; and /ʒ/ is usually pronounced as [dʒ] or [j]. The reason underlying the phenomena is the influence of Hindi. What's more, when /z/ is not word-initial, it is often pronounced as [s], and the fact is influenced by the spelling of the word.

Besides consonants, the vowels of IE also show some Indian characteristics. For example, diphthongs /eɪ/ and /əʊ/ are often pronounced as monophthongs; the BATH vowel /æ/ is mainly pronounced as [a], the LOT vowel /ɑ/ is pronounced as [ɔ], and the post-vocalic /ɹ/ tends to be silent in IE. The reasons underlying the variations are the influence of Hindi and the differences between BE and NAE. The following are summarized variation features of IE:

(1) Word-initial /p，t，k/ are not aspirated.
(2) /v/ is pronounced as [ʋ].
(3) /θ/ is pronounced as [t̪ʰ].
(4) /ð/ is pronounced as [d̪].
(6) /ʒ/ is pronounced as [ʃ] or[j].
(7) /z/ is pronounced as [s] when it is not syllable-initial.
(8) The BATH vowel /æ/ is pronounced as [a].
(9) The LOT vowel /ɑ/ is pronounced as[ɔ].
(10) /eɪ/ is pronounced as monophthong [eː].
(11) /əʊ/ is pronounced as monophthong [oː] or [ɔː].
(12) /ɹ/ after vowels are not rhotic.
(13) Function words are pronounced in weak forms.
(14) Vowels in unstressed syllables are not weak.
(15) IE pronunciation is influenced by spelling.
(16) The rules of plural-noun inflectional suffixes are:

a) When nouns or verbs end with sibilants, the inflection is added with a vowel, and pronounced as /ɪz/;

b) When nouns or verbs end with non-sibilant sounds, the inflectional suffix is /s/.

(17) The rules of past-participant inflectional suffixes are:

a) When words end with /d/ or /t/, the inflectional suffix is /ɪd/;

b) When words end with voiced sounds other than /d/, the inflectional suffix is pronounced as /d/.

c) When words end with voiceless sounds other than /t/, the inflectional suffix is pronounced as /t/ or /d/.

4. Construction of an ASR-oriented IEPD

After summarizing the variation features of IE, we begin to build an ASR-oriented IEPD mainly based on an adaptation strategy. The dictionary consists of five parts, namely the IE phoneme set, the common-word list, the Indian-word list, the acronym list and the affix list.

4.1 The IE Phoneme Set

Based on the variation features of IE and the phoneme set of NAE we design the phoneme set of IE, its IPA as well as X-SAMPA annotation being listed in Table 2 :

Consonants		Vowels	
X-SAMPA	IPA	X-SAMPA	IPA
p	p	i	i
b	b	I	ɪ
m	m	E	ɛ
f	f	{	æ
v	v	@	ə
v\	ʋ	A	ɑ
T	θ	u	u
D	ð	U	ʊ
d_d	ɖ	O	ɔ
t_d_h	ʈʰ	V	ʌ
t	t	a	a
d	d	aI	aɪ
n	n	aU	aʊ
l	l	OI	ɔɪ
r\	ɹ	3:r	ɜːr
k	k	3:	ɜː
g	g	e:	eː
dZ	dʒ	o:	oː
tS	tʃ		
S	ʃ		
Z	ʒ		
j	j		
h	h		
N	ŋ		
s	s		
z	z		
4	ɾ		
w	w		

Table 2: Phoneme Set of IE

4.2 Adaptation of CMUdict into the IE Common-word List

After comparing the differences between IE and NAE, we summarize the phonological variation features of IE, and then use the features to transform the CMUdict into the IE common-word list. As to the words with more than one pronunciations, we label them the primary pronunciation and the secondary pronunciation and note them with numbers. The following table shows some entries of the list:

Word Form	IE Pronunciation
above	@" b@U
according	@" kQr.dINg
aeroplane	" e@.ro.ple:n
although	Ql" d_do:

Table 3: Some Entries of the Common-word List

Besides the basic forms of words, the list also offers the inflectional forms of verbs (third person single, past tense, gerund), nouns (plural form), adjectives (the comparative form and superlative form) and so on. The list also includes common compounds, phrases, proper nouns, foreign words and new words listed in CMUdict.

4.3 The Indian-word List

On many occasions English lacks in its ability to convey aspects of the Indian culture in Indian people's daily communication. Some contents characteristic of the Indian local habits and culture cannot be expressed by BE or AE expressions. Only words in the Indian local languages or new words can convey the meaning. These words are usually referred to as "Indian words". Indian words are common in both the written and spoken varieties of English in India, but occur considerably more frequently in spoken Indian English than in written English taken as a whole (Balasubramanian, 2009:126, 129). These words appear frequently in Indian people's daily communication but are absent in the English pronouncing dictionaries.

We plan to collect the Indian words in three ways:1) While collecting and annotating the IE spontaneous spoken speech materials, we pick out the Indian words manually and mark their pronunciation in the Indian-word list. 2) We collect researches focusing on the Indian words, such as Dai and Gao(2012), Yan (2016), and so on, pick the Indian words out and offer their corresponding pronunciations in the Indian-word list. 3) We crawl some inflential Indian websites, especially forums and new social media (facebook, twitter, etc.) to collect spoken-style texts, extract all the words in these materials and form a word list. Then compare the word list with the CMUdict. If words in the word list are absent in CMUdict, we pick them out and put them in the Indian-word list. Then we select the Indian words manually from the new list and mark them with IE pronunciations according to the phonological variation features of IE.

4.4 The Acronym List

Since acronyms are widely used in IE, while constructing IEPD, we also try to build a acronym list. The acronyms can refer to names of places, organizations, parties, and titles of officials, etc.

4.5 The Affix List

As to the out of vocabulary (OOV) words, some researches design word/pieces mixture speech recognition systems. The main difference between the mixture speech recognition system and traditional speech recognition system lies in the fact that it can recognize OOV words. Considering that when transcribing the speech to text the ASR system will inevitably encounter OOV words, we include some widely-used prefixes and suffixes in English and offer their IE pronunciation in the affix list. If the ASR system meets new words, it can first analyze the prefixes or suffixes, then through computation infer the spelling of the words. The following table shows some affixes in affix list along with their IPA as well as X-SAMPA annotation.

Affix	**IPA**	**X-SAMPA**
anti-	æntɪ	{ntI
co-	koː	koː
-ation	eːʃən	eːS@n
-ment	mənt	m@nt

Table 4 Pronunciation of Some Affixes in IE

Till now, our team have successfully adapted CMUdict into the common-word List, while the Indian-word list, the acronym list and the affix list are still under construction.

5. Experiment

After a detailed analysis of the differences between IE and NAE pronunciation, 22 variation features of IE are summarized and formalized through comparing the pronunciation of words in the corpus and entries offered by the CMUdict. The formalized variation features thus form a variation feature set of IE phonology. Then the NAE-based CMUdict is transformed into a common-word list according to the variation feature set. Since the Indian-word list, the acronym list as well as the affix list are still under construction, we only use the common-word list in the training of IEASR.

We collect about 120 hours of IE spontaneous speech materials from programs such as "The Big Fight" , "NDTV Dialogue" from the NDTV website and annotate the speakers as well as the contents of the speech materials. Then the speech materials are cut into audio clips with the length of each audio clip being about 6-8 seconds and the sampling rate being 8K. Then from several Indian websites we crawl many texts and form a text corpus with the scale being about 1G. The audio clips are used to train the acoustic model with the Bidirectional Long and Short Time Memory (BLSTM) neural network method. The texts are used to train the language model with the 3-n gram method. Then two systems are built with the same acoustic model and language model but different pronunciation dictionaries. System 1 is trained with CMUdict, while System 2 is trained with the common-word list in IEPD. 50 audio clips of spontaneous IE speech (1148 words) are used as the test set. WERs of the two systems are 22.30% and 18.82% seperately. The result shows that System 2 outperforms System 1 with WER being 15.63% lower, as is shown in Figure 4:

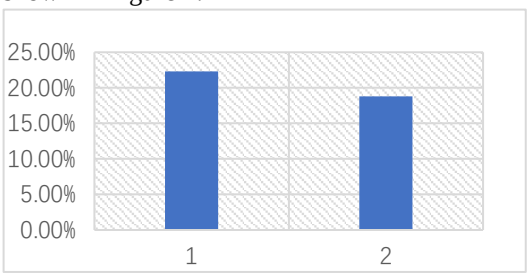

Figure 4: WERs of System 1 and System 2

From analyzing the transcribed texts of the two systems, we find that System 2 performs better in recognizing some keywords (nouns, verbs and adjectives) in the sentences, for example:

Example 1 (in each example "a" refers to manual transcription, "b" refers to transcription of System 1, and "c" refers to transcription of System 2):

test/00037a uh I don't think that we show the people have been fulfilled every time

test/00037b uh i don't think that we shove the people has in full field every time

test/00037c uh i don't think that we show the people has in fulfilled every time

Exmple 2:

test/00039a and maybe people like us need to build that ecosystem I think that's a larger role to play

test/00039b then maybe people like us need to build atico system think that's of a larger role to play

test/00039c and maybe people like us need to build that ecosystem think that's a larger role to play

In Example 1 System 2 transcribed the keywords "show" and "fullfilled" correctly while System 1 transcribed them as "shove" and "full filed". In exmaple 2 System 2 correctly transcribed"ecosystem", while System 1 transcribed it as "atico system". From the two examples we can see System 2 performs better in transcribing key words in the two sentences. In the future, we shall use more test data to further analyze the performance of the two systems.

6. Conclusion

Since IE is a relatively low-resourced variety with its local characteristics, this study selects spontaneous speech of representative Hindi speakers, analyzes and summarizes the differences between the actual pronunciation of the speakers and the entries in CMUdict, then transforms the CMUdict into an IE common-word list according to the variation features. The IE common-word list and the Indian-word list, the acronym list and affix list together form an IEPD. Since the latter three parts are not finished yet, we just use the common-word list to test whether it performs better than the AE-based CMUdict in IEASR. Through an experiment, we find that WER of the system trained with the common-word list is 15.63% lower than that trained with CMUdict when transcribing the test set. We apply phonological knowledge into constructing an ASR-oriented pronunciation dictionary for IEASR, which saves a lot of time and energy and proves a relatively good effect. In the future, we shall finish the Indian-word list, the acronym list as well as the affix list, and have more tests on the performance of IEPD with more IE spontaneous speech materials as test data.

7. Acknowledgements

This work is supported by the National Natural Science Foundation of China (No.11590771) and the Social Science Fund of Henan Province (No.2018CYY031).

8. Bibliographical References

[1] Amodei, D. , et al.(2016). Deep Speech 2 : End-to-End Speech Recognition in English and Mandarin. In *Proceedings of the 33rd International Conference on Machine Learning*. New York, NY, USA.

[2] Bansal.R.K. (1969).The intelligibility of Indian English[Z]. Monograph No.4, Hyderabad, India.

[3] Balasubramanian.2009. *Register Variation in Indian English.* John Benjamins Publishing Company, Amsterdam/Philadelphia.

[4] Dai, W.P, and Gao, L.J. (2012). Indian English: a Research on English Varieties and their Deep Mechanism. *Journal of Changchun University of Science and Technology*, (3):92-94

[5] Gimson, A.C. (1991) *English Pronunciation Dictionary*. Cambridge: Cambridge University Press.

[6] Gramley, S and Patzold, K.M. (1992). *A Survey of Modern English.* Routledge, Lndon and New York.

[7] Kachru, B.B.(1983).*The Indianization of English:the English language in India*. Oxford University Press, 1983

[8] Kachru, B.B. (1985). Standards, Codification and Sociolinguistic Realism: the English Language in the Outer Chandrika Circle. *English in the World, Teaching and Learning the Language and Literatures*, edited by Randolph Quirk and Henry Widdowson. Cambridge University Press, Cambridge and New York, pp.11-30

[9] McArthur, T. (2001). *Concise Oxford companion to the English language*. Shanghai Foreign Language Education Press.

[10] Nihalani, P, Tongue, R.K., and Hosali, P. (2005). *Indian and British English: A Handbook of Usage and Pronunciation.* Oxford University Press, Delhi.

[11] Platt, J, Weber, H, and Ho, M.L.(1984). *The New Englishes*. Routledge&Kegan Paul, London.

[12] Pull, D.K. and Kumar, G.B. (2016). Investigation of Indian English Speech Recognition using CMU Sphinx. *International Journal of Applied Engineering Research*, (6): 4167-4174.

[13] Rabiner, L and Schafer, R W. (2011). Theory and Applications of Digital Speech Processing (1st edition). Pearson, New York.

[14] Trudgill, P and Hannah, J.(2002). *International English: a Guide to Varieties of Standard English* (4th edition). Arnold, London

[15] Wells, J.C. (2005). *Longman Pronunciation Dictionary*. The Commercial Press, Beijing.

[16] Yan, Y.Y. On Indian Military Abbreviations and Acronyms and their Translation. *Journal of Asian and African Studies*, (1):139-150

9. Language Resource References

Kingsbury, Paul, et al. CALLHOME American English Lexicon (PRONLEX) LDC97L20. Web Download. Philadelphia: Linguistic Data Consortium, 1994.

Baayen, R H., R Piepenbrock, and L Gulikers. CELEX2 LDC96L14. Web Download. Philadelphia: Linguistic Data Consortium, 1995.

Canavan, Alexandra, David Graff, and George Zipperlen. CALLHOME American English Speech LDC97S42. Web Download. Philadelphia: Linguistic Data Consortium, 1997.