Research Inventy: International Journal of Engineering And Science
Vol.4, Issue 7 (July 2014), PP 06-11
Issn (e): 2278-4721, Issn (p):2319-6483, www.researchinventy.com

Phonetic Transcription- A Framework for Phonetic Representation of Sound Structures

Dr. M Hanumanthappa[1], Rashmi S[2], Jyothi N M[3]
[1]*Associate Professor, Department of Computer Science and Applications, Bangalore University, Bangalore-56.*
[2]*Research Scholar, Department of Computer Science and Applications, Bangalore University, Bangalore-56.*
[3]*Assistant Professor, Department of MCA, Bapuji Institute of Engineering and Technology, Davangere.*

ABSTRACT: *Implementing phonetics to Natural language is a Herculean task. The first step for the phonetic implementation is to generate a phonetic dictionary. Phonetic dictionary is very important as it plays a vital role in identifying the essential components of the gigantic vocabulary in the speech recognition system of natural language. Language like English needs phonetic transcriptions because the English spelling does not tell us how to pronounce it. Pronunciation is very important for communication as it vitalizes rapid transition to practical concern. There is lot of ongoing research in multilingual speech recognition. Phonetic dictionary is utmost important for any speech recognition system and hence the key goal of the paper is to build a phonetic dictionary for English language. The paper aims at building dictionary for identifying the phonetic transcriptions of International Phonetic Alphabet (IPA) and American English alphabets phonetically. The paper also gives a detailed explanation about the rules of the phonemes.*

KEYWORDS*: Natural Language Processing (NLP), Phonetics, Acoustic phonetics, Phonemes, Phonetic transcriptions, WordNet.*

I. INTRODUCTION

According to the WordNet the phonetics is defined as a branch of NLP which is concerned with the speech production and perception and a detailed analysis of acoustics. Acoustic phonetics is a branch of phonetics which deals with the study of sounds that is made by the vocal organs of the human in order to produce the sound. The acoustic knowledge is very important to locate the underlying phonetic representation. Acoustics discusses how the sound travels from speakers' mouth to listeners' ears. The challenges of identifying the language and storing the corresponding sound for phonetic transcription are tedious and tiresome. The International Phonetic Alphabet (IPA) and the acoustic modeling of various phonemes have shown rationally good improvement for the implementation and understanding of the sound structure. Various approaches such as Statistical Machine Translation (SMT), language identification system (LID) merges gap between the encoding and the decoding process by using multi stream approach [1].

Phonemes are the basic building block for the language which is combined with the other phoneme to build a meaningful unit such as words and morphemes. Phonemes are the individual or group of sound units. The phonetic codes include syllable code, syllable time and syllable rhythm. The phonemes form the bedrock for initial and final code form to represent the phonetic structure with sensibility and sub-information of linguistics in syllable code [2].

Phonetic transcriptions tell us how the word has to be pronounced. These phonetic transcriptions are written in IPA where each and every English alphabet has its own symbol. For example the IPA based phonetic transcription for the words such as *no* is noʊ, and the transcription of *do* is duː. Though both of these words end with the same letter they have different sound and the phonetic transcriptions are different. The transcription for the word *alphabet* is [aɪ pʰiː eɪ]. Phonetic transcriptions are helpful because there are no convincing empirical results which help us to sort the pronunciation training [5]. The below table mention some of the examples for phonetic transcriptions.

main stress	/ˌek.spekˈteɪ.ʃən/ expectation
secondary stress	/ˌriːˈtell/ retell
syllable division	/ˈsɪs.təm/ system

Table-1 phonetic transcriptions

II. LITERATURE SURVEY

It is said that there are around 4000 languages spoken today across the world and most of these languages have limited linguistic knowledge and speech data/resources are available. Author Akbacak, M et al have proposed series of speech retrieval algorithms in order to leverage the already existing algorithms. The algorithms employ confusion-embedded hybrid pronunciation networks, and lattice-based phonetic search within a proper name retrieval task. Latin-American Spanish are used as the target language. After searching for queries consisting of Spanish proper names in Spanish Broadcast News data, we demonstrate that retrieval performance degradations (due to data sparseness during automatic speech recognition (ASR) deployment in the target language) are compensated by employing English acoustic models [9]. Speech recognition system for various languages have been studied and developed. Arabic speech recognition system is developed at Cambridge University by Gales M J F et al [6]. Author has shown a simple scheme for automatically generating associations among diverse pronunciations for use in training and reducing the phonetic out-of-vocabulary rate. Author Palanisamy K et al, has developed Tamil pronunciation dictionary by incorporating the visual actions of the organs for improving communication skills [7]. Phonetic transcriptions by using Speech Assessment Method Phonetic Alphabet (SAMPA) for Romanian language has been proposed by Domokos, J et al. Development of phonetic dictionary and also the system architecture is explained in the paper [8].

In the next section various methodologies for building the phonetic dictionary and the associated rules are discussed.

III. METHODOLOGIES

Today almost all the English dictionaries have audio recordings. So even with this, the need of phonetic transcriptions is prominent because of the following reasons [3]

- To have a good communication in English first we have learn the pronunciations and English language is polysemy in nature. So learning the pronunciations through sound is apt in these kinds of situations.

- When we are listening to the pronunciations we might not be sure whether we heard ʊ or ə, ɒ or ʌ, s or z, etc., this is due to the lack of experience and the quality of the particular sound is poor. Hence reading phonetic transcriptions will make the phonemes clear as we will see the sound symbols of all the alphabets in a word

- Dictionaries often show multiple transcriptions but have the same pronunciations so the better way to find the reasons for this ambiguity would be to study the information about the particular transcriptions.

- The computer which we are using might not have the speakers and hence the audio recording of the particular sound cannot be heard. Sometimes even if the speakers are available we do not want to disturb the public with the sound. In these situations transcriptions will greatly help.

Word Stress: A word is combination of syllables. In every word, one or more letters are pronounced strongly. This is called word stress. The dictionary which is shown in this paper also contains the word stress. Pronunciation for English language is broadly classified into American and British transcriptions. In American English, the sound of **r** is always stressed where as in British English **r** is not stressed. Example: arm/ father are pronounced differently in both of the accent.

IPA: International Phonetic Alphabet is the abbreviation for IPA. It defines the standard phonetic symbol for every alphabet in English language. The IPA symbols are usually written in the Latin symbols. IPA defines the standard sound representation for oral language. IPA is considered as the standard for linguistics. However there is also the American phonetic alphabet because of the difficulties faced with IPA such as, it is tough to type the IPA symbols with the normal keyboard. IPA also increases the error rate by withholding the awkwardness reading the transcriptions which are hand written.

IV. BUILDING PHONETIC DICTIONARY

Phonetic dictionary is not as simple as it looks because applying sounding mechanisms to make computer auto detect the word and give the pronunciation is difficult. There are various rules that need to be followed while building the phonetic transcriptions. The rules are [3]

[1]. Almost all dictionaries use the e symbol for the vowel in *bed*. The problem with this convention is that e in the IPA does *not* stand for the vowel in *bed*; it stands for a different vowel that is heard, for example, in the German word *See*<u>*l*</u>*e*. The "proper" symbol for the *bed* vowel is ɛ (do not confuse with ɜ:). The same goes for eə vs. ɛə.

[2]. In ər and ɜːr, the ʳ is not pronounced in BrE, unless the sound comes before a vowel (as in *answering, answer it*). In AmE, the ʳ is always pronounced, and the sounds are sometimes written as ɚ and ɝ.

[3]. In AmE, ɑː and ɒ are one vowel, so *calm* and *cot* have the same vowel. In American transcriptions, *hot* is written as hɑːt.

[4]. About 40% of Americans pronounce ɔː the same way as ɑː, so that *caught* and *cot* have the same vowel. See *cot-caught merger*.

[5]. In American transcriptions, ɔː is often written as ɒː (e.g. *law* = lɒː), unless it is followed by r, in which case it remains an ɔːr.

[6]. In British transcriptions, oʊ is usually represented as əʊ. For some BrE speakers, oʊ is more appropriate (they use a rounded vowel) — for others, the proper symbol is əʊ. For American speakers, oʊ is usually more accurate.

[7]. In eəʳ ɪəʳ ʊəʳ, the r is not pronounced in BrE, unless the sound comes before a vowel (as in *dearest, dear Ann*). In AmE, the r is always pronounced, and the sounds are often written as er ɪr ʊr.

[8]. All dictionaries use the r symbol for the first sound in *red*. The problem with this convention is that r in the IPA does *not* stand for the British or American *r*; it stands for the "hard" *r* that is heard, for example, in the Spanish word *rey* or Italian *vero*. The "proper" symbol for the *red* consonant is ɹ.

[9]. In American English, t is often pronounced as a *flap t*, which sounds like d or (more accurately) like the quick, hard *r* heard e.g. in the Spanish word *pero*. For example: *letter*. Some dictionaries use the t˔ symbol for the *flap t*.

Table II describes the English alphabets and the corresponding IPA and American phonetic alphabets. Word stress is also highlighted.

Table II IPA and phonetic Transcriptions

English Alphabet	a	b	c	d	e	f	g	h	i	J	k	l	M	n	o
IPA/American Phonetic Alphabet	ə	bi	si	di	I	ɛf	ǰi/dʒi	etʃ/eč	aj/ay	dʒe/ǰe	ke	ɛl	ɛm	ɛn	o
Stress Marked across the Alphabets	é	bí	sí	dí	í	ɛ́f	ǰí/dʒí	étʃ/éč	áj/áy	dʒé/ǰé	ké	ɛ́l	ɛ́m	ɛ́n	ó

English Alphabet	p	Q	r	s	t	u	v	W	x	Y	z
IPA/American Phonetic Alphabet	pi	kyu/kju	ɑr	ɛs	ti	ju/yu	vi	dəbəlju/dəbəlyu	ɛks	waj/way	zi
Stress Marked across the Alphabets	pí	kyú/kjú	ár	ɛ́s	tí	jú/yú	ví	də́bəlju/də́bəlyu	ɛ́ks	wáj/wáy	zí

Table III shows the IPA phonetic symbols and its phonetic description. The short cut key defines the letters that we have to type in order to get the appropriate phonetic symbol.

Alphabets	**Phonetic symbol**	**Phonetic description for the symbol**	**Short cut key**
	ɑ	open back unrounded vowel	(Ctrl+A)
	æ	near-open front unrounded vowel	(Ctrl+AA)
	ɐ	near-open central vowel	(Ctrl+AAA)
A	ɑ̃	nasalized open back unrounded vowel	(Ctrl+AAAA)
	β	voiced bilabial fricative	(Ctrl+B)
	ɓ	voiced bilabial implosive	(Ctrl+BB)
B	ʙ	bilabial trill	(Ctrl+BBB)

Alphabets	**Phonetic symbol**	**Phonetic description for the symbol**	**Short cut key**
M	ɱ	labiodental nasal	(Ctrl+M)
	ŋ	velar nasal	(Ctrl+N)
	ɲ	palatal nasal	(Ctrl+NN)
	ɴ	uvular nasal	(Ctrl+NNN)
N	ɳ	retroflex nasal	(Ctrl+NNNN)
	ɔ	open-mid back rounded vowel	(Ctrl+O)
O	œ	open-mid front rounded vowel	(Ctrl+OO)

	ɕ	voiceless alveo-palatal fricative	(Ctrl+C)
C	ç	voiceless palatal fricative	(Ctrl+CC)
	ð	voiced dental fricative	(Ctrl+D)
	d͡ʒ	voiced postalveolar affricate	(Ctrl+DD)
	ɖ	voiced retroflex plosive	(Ctrl+DDD)
D	ɗ	voiced alveolar implosive	(Ctrl+DDDD)
	ə	mid-central vowel	(Ctrl+E)
	ɚ	rhotacized mid-central vowel	(Ctrl+EE)
	ɵ	close-mid central rounded vowel	(Ctrl+EEE)
E	ɘ	close-mid central unrounded vowel	(Ctrl+EEEE)
	ɛ	open-mid front unrounded vowel	(Ctrl+3)
	ɜ	open-mid central unrounded vowel	(Ctrl+33)
	ɝ	rhotacized open-mid central unrounded vowel	(Ctrl+333)
	ɛ̃	nasalized open-mid front unrounded vowel	(Ctrl+3333)
F	ɞ	open-mid central rounded vowel	(Ctrl+33333)
	ɠ	voiced velar implosive	(Ctrl+G)
	ɢ	voiced uvular plosive	(Ctrl+GG)
G	ɢ̓	voiced uvular implosive	(Ctrl+GGG)

	ø	close-mid front rounded vowel	(Ctrl+OOO)
	ɒ	open back rounded vowel	(Ctrl+OOOO)
	ɔ̃	nasalized open-mid back rounded vowel	(Ctrl+OOOOO)
	œ	open front rounded vowel	(Ctrl+OOOOOO)
P	ɸ	voiceless bilabial fricative	(Ctrl+P)
	ɾ	alveolar tap	(Ctrl+R)
	ʁ	voiced uvular fricative	(Ctrl+RR)
	ɹ	alveolar approximant	(Ctrl+RRR)
	ɻ	retroflex approximant	(Ctrl+RRRR)
	ʀ	uvular trill	(Ctrl+RRRRR)
	ɽ	retroflex flap	(Ctrl+RRRRRR)
R	ɭ	alveolar lateral flap	(Ctrl+RRRRRRR)
	ʃ	voiceless postalveolar fricative	(Ctrl+S)
S	ʂ	voiceless retroflex fricative	(Ctrl+SS)
	θ	voiceless dental fricative	(Ctrl+T)
	t͡ʃ	voiceless postalveolar affricate	(Ctrl+TT)
	t͡s	voiceless alveolar affricate	(Ctrl+TTT)
T	ʈ	voiceless retroflex plosive	(Ctrl+TTTT)

	ɥ	labial-palatal approximant	(Ctrl+H)
	ɦ	voiced glottal fricative	(Ctrl+HH)
	ħ	voiceless pharyngeal fricative	(Ctrl+HHH)
	ɧ	Sje-sound	(Ctrl+HHHH)
H	ʜ	voiceless epiglottal fricative	(Ctrl+HHHHH)
	ɪ	near-close near-front unrounded vowel	(Ctrl+I)
	ɨ	near-close central unrounded vowel	(Ctrl+II)
I	ɨ	close central unrounded vowel	(Ctrl+III)
	ʝ	voiced palatal fricative	(Ctrl+J)
	ɟ	voiced palatal plosive	(Ctrl+JJ)
J	ʄ	voiced palatal implosive	(Ctrl+JJJ)
	ɬ	velarized alveolar lateral approximant	(Ctrl+L)
	ɬ	voiceless alveolar lateral fricative	(Ctrl+LL)
	ʟ	velar lateral approximant	(Ctrl+LLL)
	ɭ	retroflex lateral approximant	(Ctrl+LLLL)
L	ɮ	voiced alveolar lateral fricative	(Ctrl+LLLLL)
	ʊ	near-close near-back rounded vowel	(Ctrl+U)
	ʊ	near-close central rounded vowel	(Ctrl+UU)
U	ʉ	close central rounded vowel	(Ctrl+UUU)
	ʌ	open-mid back unrounded vowel	(Ctrl+V)
	ʋ	labiodental approximant	(Ctrl+VV)
V	ⱱ	labiodental flap	(Ctrl+VVV)
	ʍ	voiceless labio-velar approximant	(Ctrl+W)
	ɯ	close back unrounded vowel	(Ctrl+WW)
W	ɰ	velar approximant	(Ctrl+WWW)
X	χ	voiceless uvular fricative	(Ctrl+X)
	ʎ	palatal lateral approximant	(Ctrl+Y)
	ɣ	voiced velar fricative	(Ctrl+YY)
	ʏ	near-close near-front rounded vowel	(Ctrl+YYY)
Y	ɤ	close-mid back unrounded vowel	(Ctrl+YYYY)
	ʒ	voiced postalveolar fricative	(Ctrl+Z)
	ʐ	voiced retroflex fricative	(Ctrl+ZZ)
Z	ʑ	voiced alveolo-palatal fricative	(Ctrl+ZZZ)

Table III Phonemes

V. CONCLUSION

Pronunciation is very important in today's communication and currently there has been a shift from linguistic competencies to a broader level of communicative compliances. Effective communication is always based on the good pronunciation. Pronunciation is reckoned as not just production of the right phonemes. It forms the foundation for the next level of speech analysis. With adequate pronunciation skill one can fly to a new horizon by achieving professional responsibility. But with the increasing ambiguities in the natural language it is difficult to judge the right pronunciation. The goal of this paper was to provide a phonetic dictionary. This forms the root for the later part of the research work which is to build the interface for recognizing phonetic structure of sounds generated by the natural language mostly in English. The dictionary shown in the paper is for Latin-American English alphabets only. It is intentionally kept for these languages in order to limit the resources. The same rules will form the ground work to recognize the sound of other native languages such as Kannada, Telugu and others. Various researchers can get benefit from the paper by looking into the rules and the dictionary for building the sound recognition system. Future scope of the work will be on developing a tool for phonetic dictionary by applying the rules specified in the paper.

REFERENCES:

[1]. *A first speech recognition system for Mandarin-English code-switch conversational speech* Ngoc Thang Vu ; Dau-Cheng Lyu ; Weiner, J. ; Telaar, D. ; Schlippe, T. ; Blaicher, F. ; Eng-Siong Chng ; chultz, T. ; Haizhou Li , Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on Digital Object Identifier: 10.1109/ICASSP.2012.6289015 Publication Year: 2012 , Page(s): 4889- 4892

[2]. *Notice of Retraction A kind of Chinese language Phonetic input output system code*, Lu Qiao ; Wan Pu ; Zhang Li ; Zhu Dao-yong Computer Science and Information Technology (ICCSIT), 2010 3rd IEEE International Conference on Volume:9 Digital Object Identifier: 10.1109/ICCSIT.2010.5563773 Publication Year: 2010 , Page(s): 508- 511

[3]. *Tomasz P. Szynalski*, Proceedings of the International Research Conference "*Metamorphoses of the Mind: Out of One's Mind, Losing One's Mind, Lightmindedness, Mindlessness*" Organized by the Latvian Academy of Culture and Goethe Institut Riga Riga, 21–23 September 2006 *http://www.antimoon.com/how/pronunc-trans.htm*

[4]. http://www.antimoon.com/how/pronunc-soundsipa.htm#chart

[5]. Marcus Otlowski, *Pronunciation: What Are the Expectations?*, The Internet TESL Journal, Vol. IV, No. 1

[6]. *Development of a phonetic system for large vocabulary Arabic speech recognition*, Gales, M.J.F. ; Cambridge Univ., published in Automatic Speech Recognition & Understanding, 2007. ASRU. IEEE Workshop on Date of Conference:9-13 Dec. 2007 Page(s):24 – 29 E-ISBN :978-1-4244-1746-9 Print ISBN: 978-1-4244-1746-9 INSPEC Accession Number: 9832782

[7]. *Audio visual based pronunciation dictionary for Indian languages.* Palanisamy, K. *;* Published in:Technology for Education (T4E), 2010 International Conference on Date of Conference: 1-3 July 2010 Page(s): 82 – 84 E-ISBN : 978-1-4244-7361-8 Print ISBN:978-1-4244-7362-5 NSPEC Accession Number: 11476081

[8]. *100K+ words, machine-readable, pronunciation dictionary for the Romanian language*, Domokos, J. ; Buza, O. ; Toderean, G. Signal Processing Conference (EUSIPCO) 2012 IEEE Proceedings of the 20th European Date of Conference:27-31 Aug. 2012 Page(s):320 – 324 ISSN :2219-5491 Print ISBN:978-1-4673-068-0 INSPEC Accession Number:13072796

[9]. *Spoken Proper Name Retrieval for Limited ResourceLanguages Using Multilingual Hybrid Representations*, Akbacak, M. ; Hansen, J.H.L. Audio, Speech, and Language Processing, IEEE Transactions on Volume: 18 , Issue: 6 Digital Object Identifier: 10.1109 /TASL .2009 .2035785 Publication Year: 2010 , Page(s): 1486 - 1495