_____

A PROPORTIONAL STUDY OF UK AND US ENGLISH ACCENTS IN RECOGNITION AND SYNTHESIS

V.Bujji Babu
Asst.Prof of English
DVR&HS MIC College of Technology
Kanchikacherla, Andhra Pradesh, India

_____

Abstract:

We have a tendency to gift a comparative study of the accents options of two major English accents: British English and English. Experiments examined the deterioration in speech recognition ensuing from the couple between English accents of the input speech and therefore the speech models. Mismatch in accents will increase the error rates by quite a hundred per cent. Hence an in-depth study of the acoustic correlates of accent mistreatment modulation and pitch characteristics was done . Accents variations area unit acoustic manifestations of variations in the period, pitch and intonation pattern and of course the differences in phonetic transcriptions. Particularly, British speakers possess a lot of vessel pitch rise and fall pattern and lower average dig in most of the vowels. Finally a doable suggests that to convert English accents is usually recommended supported higher than analysis.

Keywords: study, accent, British, American etc

_____

1. Introduction:

Two of the foremost vital remaining obstacles to reliable high-performance speech recognition systems area unit noise and speaker variations. An important aspect of speaker variation is accepted. However, current speech recognizers are a unit trained on a particular national accent cluster (e.g.UK or North American nation English accents), and should have a major deterioration in performance once process accents unseen within the coaching information. In [3] J.C. Wells delineated the term accent as a pattern of pronunciation utilized by a speaker for whom English is that the language or additional typically, by the community or social grouping to which he or she belongs.

There has been significant analysis conducted on understanding the causes and also the acoustics correlates of native English accent. A study in [3] examined a range of native English accents from a linguistics purpose of reading. Recently additional cantered studies are created on acoustic characteristics of English accents. In [4] a technique is represented to decrease the popularity error rate by mechanically generating the accent lexicon through comparison of normal transcription with decoded phone sequence. In [1], instead of victimization phonetic symbols, different regional accents are synthesized by an accent- independent keyword lexicon. During synthesis, the input text is **1st** transcribed as keyword lexicon.

Until post-lexical processes, accent dependent phoneme rules were applied to subsume such options as /t//d/ topping in US English, or r-linking in British English. The advantage of this methodology is that it avoids applying totally different phonetic symbols to represent varied accents. In addition, [2] established a voice conversion system between British and US English accents by HMM-based spectral mapping with set rules for mapping 2 completely different phone sets. However, it still has some residual of original supply accent characteristics within the regenerate result.

In this paper, experiments began with cross accent recognition to quantify the accent effects between British accent (BrA) and American accent (GenAm) on speech recognition. A further elaborate acoustics feature study of English accent victimization period, intonation and frequency characteristics was performed.

2. Cross Accent Recognition:

At first, a group of easy experiments was doled out to quantify the impact of accents on the speech recognisers with accent specific dictionaries. The model coaching and recogniser used here square measure supported HTK [9].British accent speech recogniser was trained on Continuous Speech Recognition Corpus (WSJCAM0). American accent speech recogniser was trained on WSJ. Test sets used are WSJ si_dt_05 si_et_05 and

_____

WSJCAM si_dt5b, each containing 5k words. Both recognisers employ 3-state left-to-right HMMs.The options utilized in experiments were thirty-nine MFCCs with energy and their differentiation and acceleration.

Accent	British model	American model
British input	12.8	29.3
American input	30.6	8.8
Average	21.7	19.1

Table 1: the last word error rate of cross accents speech recognition between British and yank accent Table one shows that for this information the American English achieves thirty-first Mismatched accent of the speaker and therefore the recognition system deteriorates the performance. The result was obtaining worse by 139% for recognizing British English with yank models and 232% for recognizing the English language with British models. The results square measure supported word models compiled from triphone HMMs with three states per model and twenty mixtures Gaussians per state. The next section examines the acoustics options of each English accent in a shot to spot wherever the most distinction lies additionally to the variation in pronunciation.

3. Analysis of Acoustic Options Folks and UK English Accents Duration:

Figure one shows that the vowel durations at the beginning and therefore the finish of sentences in brassiere is shorter than that in GenAm. This could result in the subsequent reason. However, Americans tend to appreciate a lot of acoustically complete pronunciation. Table two offers the comparison of two information in speaking rate. The speaker rate of Wsjcam0 is seven.8% higher than that of WSJ.This is in accordance with comparison in phone period in Figure one. The results of these comparisons are shown in Figure 1.Note that results in area unit solely bestowed for models common to each system phones sets.

Speak Rate (no/sec)	Phone	Word
Wsjcam0	9.77	3.04
Wsj	10.39	2.82

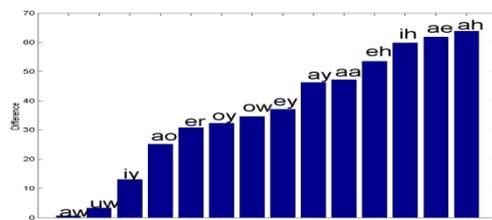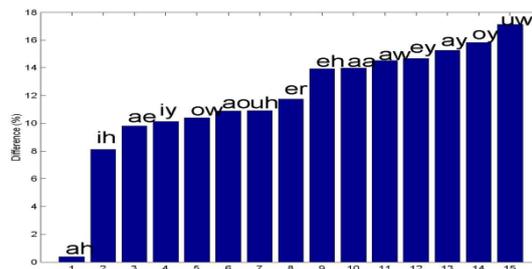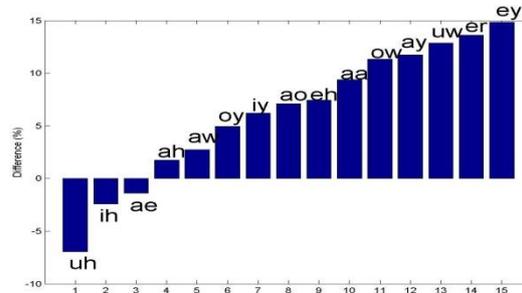Table 2: Speak rate in Phone and word from Wsjcam0 and WSJ

Figure 1: Difference of Vowel duration of Gen Am and Br A at the utterance starts and ends Pitch Characteristics. Table three and four list average pitch values and numbers of speakers from each database. Figure a pair of displays the distinction of average vowel pitch frequency of male speakers of 2 accents whereas Figure three shows the corresponding comparison of feminine speakers. Even undergarment has a lower average pitch than GenAm within the whole phone set, for the common vowels, their average eat undergarment is still rather more under corresponding part in GenAm.It is attention-grabbing to notice that for many of vowels, British speakers give lower pitch than American counterparts. For British female speakers, it's 118% lower than American female on average while it drops down to 7.7% when compared with British male and American male in the common set vowels. Besides, it is often noted that British male speakers gain high average eat 3 vowels: uh, ih and ae.

Speaker No.	Male	Female
Wsjcam0	112	93
WSJ	37	41

Table 3: Number of speakers Wsjcam0 and WSJ

Avg Pitch	Male	Female
Wsjcam0	115.8 Hz	196.2 Hz
WSJ	127.8 Hz	208.9 Hz
Difference	9.4 %	5.7%

Table 4: Average pitch of Wsjcam0 and WSJ

Oray's Publications

Impact Factor: 4.845(SJIF) **Research Journal Of English (RJOE)** Vol-4, Issue-1, 2019

www.rjoe.org.in   An International Peer-Reviewed English Journal   ISSN: 2456-2696

Indexed in: International Citation Indexing (ICI), International Scientific Indexing (ISI), Cosmos, Directory of Research Journal Indexing (DRJI) & Cosmos.
_____

Prosody is typically created from Intonation-groups, Pitch Event and accent. Intonation-groups square measure composed of a sequence of pitch events at intervals phrase. Pitch Event could be a combination of a pitch rise and fall. In [6], an increase fall affiliation (RFC) model was applied to model the prosody by Legendre polynomial perform [a1, a2, a3], where a1, a2, a3, referred to as separate Legendre Polynomial Coefficients, were associated with the typical contour, average contour slope and the average trend of the slope within that pitch accent. Rise and fall are detected in step with f0 contour. Based on this, experiments were created on computing the typical pattern of pitch accents(Fall and Rise solely during this case) to explore the numerical distinction of each accent in intonation. Figure 4(a) illustrates the type of rising and fall patterns from each male and feminine speakers. It is noticeable that British speakers shall have vessel rise and fall than yank speakers. In addition, it's additionally noticeable that pitch very narrows towards the top of associate degree vocalization as [8]. Further to the results that yank speaker tends to talk lower in final words of sentences. Figure 4(b) indicates that bandeau Rise pattern within the last words is way a lot of vessel than that of GenAm with pitch amendment rate of forty-eighth and thirty seconds severally. In distinction,

the autumn pattern is nearly the same in either figure. Then British speakers possess abundant vessel accent than yank speakers.

5. Discussions and Conclusion:

We have conferred an in-depth study of acoustic options regarding major English accents: bandeau and Gen Am. In addition to the numerous distinction in acoustics, the slope of Rising and Fall accent also exhibits great difference. British speakers tend to talk with lower pitch however higher pitch modification rate, especially in the rising accent. Future experiments are to be extended to alternative context- dependent pitch pattern analysis besides auditory communication finish. In general, accent conversion/synthesis may well be simplified into two aspects: acoustics and acoustics. Beep lexicon and CMU lexicon expressly show the acoustics distinction between 2 accents in terms of phone substitute, delete and insert. Therefore, the accent synthesis is planned to hold on by 2 steps for future experiments.

1) Pronunciation modeling by transcribing GenAm by bandeau phones to map phonetic distinction of 2 accents [4] or vice verse.

2) Prosody modification [7] [8].

By applying the Tilt model base on decision-tree HMM, tilt parameters square measure modified per higher than analysis. The advantage of the Tilt model lies in its continuous tilt parameters, that higher describe the intonation than RFC models or FUJISAKI models [7]. New prosody is then synthesized once dynamic tilt parameters according to higher than study

References:

- Susan Fitt, Stephen Isard, Synthesis of Regional English Using A Keyword Lexicon. Proceedings Euro speech 99, Vol. 2, pp. 823-6.
- Ching-Hsiang Ho, Saeed Vaseghi, Aimin Chen, Voice Conversion between UK and US Accented English, Euro speech 99.
- J.C. Wells, Accents of English, volume:1,2, Cambridge University Press, 1982
- Jason John Humphries, Accent Modeling and Adaptation in Automatic Speech recognition, PhD Thesis, Cambridge University Engineering Department
- Alan Cruttenden, Intonation, Second Edition 1997
- Ching-Hsiang Ho, Speaker Modeling for Voice Conversion, PhD thesis, Department of Computing and Electronic Engineering, Brunel University
- Thierry Dutoit, Introduction to text-to-speech synthesis, Kluwer (1997)
- Paul Taylor, Analysis and Synthesis of Intonation using Tilt Model, Journal of the Acoustical Society of America. Vol 107 3, pp. 1697-1714.