THE ROLE OF SEGMENTAL INFORMATION IN SPEAKER DIALECT IDENTIFICATION

NAVYA KESAV

17SLP027

A Dissertation Submitted in Part Fulfillment of Degree of

Master of Science (Speech- Language Pathology)

University of Mysore, Mysuru

ALL INDIA INSTITUTE OF SPEECH AND HEARING

MANASAGANGOTHRI,

MYSORE- 570006

MAY 2019

CERTIFICATE

This is to certify that this dissertation entitled ***"The role of segmental information in speaker dialect identification"*** is a bonafide work submitted in part fulfilment for degree of Master of Science (Speech-Language Pathology) of the student (Registration Number: 17SLP027). This has been carried out under the guidance of a faculty of this institute and has not been submitted earlier to any other University for the award of any other Diploma or Degree.

Dr. M. Pushpavathi

Director

All India Institute of Speech and Hearing

Manasagangotri, Mysuru- 570006

Mysuru,

May, 2019

CERTIFICATE

This is to certify that this dissertation entitled ***"The role of segmental information in speaker dialect identification"*** is a bonafide work submitted in part fulfilment for degree of Master of Science (Speech-Language Pathology) of the student (Registration Number: 17SLP027). This has been carried out under my supervision and guidance. It is also been certified that this dissertation has not been submitted earlier to any other University for the award of any other Diploma or Degree.

Mysuru

May, 2019

Dr. T. Jayakumar

Guide

Associate Professor in Speech Sciences

Department of Speech Language Sciences

All India Institute of Speech and Hearing

Manasagangotri, Mysuru- 570006

DECLARATION

This is to certify that this dissertation entitled ***"The role of segmental information in speaker dialect identification"*** is the result of my own study under the guidance of Dr. T. Jayakumar, Reader in Speech Sciences, Department of Speech- Language Sciences, All India Institute of Speech and Hearing, Mysuru, and has not been submitted earlier to any other university for the award of any other Diploma or Degree.

Mysuru,
May, 2019

Registration No. 17SLP027

TABLE OF CONTENT

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENT

I would like to thank **Amma, Acha** and **Amru** for their constant support all through these years and providing me with whatever best they could get and for being role models in my life.

I would like to thank **Jayakumar Sir** for his relaxed thoughful insight wihout whom the learning curve would have been very much steeper. I thank you sincerely for spending your valuable time and effort in guiding me to complete this study.

Vijayakumar uncle, Shobhi aunty and **Namratha**, thank you so much as this study would have not been possible without your contribution. Big thanks to all the participants for their timely cooperation and patience throughout the study.

I pour out all my love and thanks to my **grandparents** for helping me out with the language for the study.

Santhosha Sir, thank you for the statistical support without which the study would have been meaningless.

Malavi, Reny, Sruthi, Anagha and **Rishi,** my the bests, love you for the indefinite support and motivation you guys have always provided me and for always standing beside me during the happy and hard movements.

Jes bro, thanks for helping me always to clear my confusions and narrow down to specific things. Your constant nagging and guidance has helped me to complete this study.

Reuben sir, Vasu akka, Priya akka and **Yaseen sir,** thank you for your valuable input and helping me survive in this by providing timely guidance whenever necessary.

Sreedevi Ma'am, Yeshoda Ma'am and **Jayashree Ma'am,** thank you so much for giving me your timely expertise by expanding my thoughts and pointing out the various dimensions I could work upon for the study.

Thereotical base is nothing without appropriate application in clinical situations. Thank you so much **Sujatha Ma'am, Vijayashree Ma'am, Swapna Ma'am, Goswami Sir, Rajasudhakar Sir, Gopishankar Sir, Gopikishore Sir, Gayathri Ma'am** for helping with this.

Safa, Keren, Pratyasha and **Dafiah,** I cannot imagine hostel life without you guys. You guys made it fun throughout. I have no words to thank my **classmates** for giving me the best of times in these two years. Honestly I have never been in a class so amazing.

Renita and **Farzeena,** my dissertation partners, love you for being supportive throughout these years.

Teenu, Shreya and **Swathi,** love you guys. You were amazing.

My dear teachers at AIMS; **Maria Ma'am,** because of whose inspiration I am here, thanks a ton Ma'am. **Prem Sir, Maya Ma'am,** thanking you for teaching me how to survive in all kinds of places. **Meera Ma'am** and **Mythili Ma'am,** I feel lucky to be your student and thank you, it is because of you my interest for dysphagia has grown to bigger heights. **Rinu, Sminu, Sradha, Neha, Manju,** love you guys for making my life at AIMS livable.

Ashi and **Racheal,** girls thanks a lot for helping me in different ways.

Finally, **NIMHANS** and **COMDEALL** faculties, special thank you to them for making me a strong clinician and making me realise my potentials.

Thank you all !

Chapter 1

INTRODUCTION

Language is defined as a "purely human and non-instinctive method of communicating ideas, emotions and desires by means of a system of voluntarily produced symbols (Sapir, 1921). It has been described as a social possession besides an individual possession (Kumari, 2009). Further language has also been defined as a "single linguistic norm" or to a "group of related norms" and dialect as "one of the norms". A dialect is constituted by "different varieties of a single language".

Dialect is a "regionally specific subdivision of a language" (Robert, 2006). Dialect differs socially, geographically and structurally. Variations have been accounted in dialects with respect to phonetics, phonology, morphology, syntax and lexicon (Petyt, 1980). Studies exist wherein it has been investigated how the perception of /l/ and /n/ in English varied across different dialects of Mandarin (Li, 1951). Differences have been reported in diphthong perception in Binghamton, Northern United states, Birmingham, Southern United states dialects (Makashay, 1999). Differences in vowel production occur in six regional varieties of American English (Clopper, Jong and Pisoni, 2005). There have been studies documenting linguistic variations in phonological rules, subject-verb agreement and nominal plural marking in Brazilian Portuguese dialect (Guy, 1981). An Indian study has been reported which showed the difference in phonological system in Awadhi, Bagheli, Bhojpuri, Bundeli, Haryanvi, Kanauji and Khari Boli dialects of Hindi (Mishra & Bali, 2011). Difference in stress patterns in the Nalbaria dialect of Assamese language have also been reported (Patgiri, 2011). The phonological and morphological differences found in the Northern, Central and Southern Malayalam are variations in personal pronouns, particles, morphemic replacements, phonemic replacements (Subramaniam (ed) 1993). He has also reported that Malayalam dialects can be distinguished based on suprasegmentals features. They can vary with intonation and pauses between words also. A recent study reported the differences in production of retroflex sound in the Muslim dialect of Malayalam. In the central Malabar Muslim dialect the voiced retroflex approximant sound /ɻ/ is substituted with the voiced palatal approximant /j/ even though the two sounds are produced as two different phonemes in other Malayalam dialect (Jayaraj, 2017).

Several researchers have reported significant differences in acoustic parameters across dialects (Jacewicz, Fox, and Salmons, 2007; Chládková, 2011; Abramson & Ren, 1990). Differences have been observed in vowel duration in conjunction with the spectral information in Swiss French versus Parisian French (Miller, Mondini, Grosjean & Dommergues, 2011). A similar study was done recently in the standard Nepali language versus three varieties of Nepali (Lohagun, 2018). The standard variety Nepali speakers had longer vowel duration when compared to the other three variants (Sikkim and speakers from the Darjeeling and Alipurduar districts of West Bengal). Acoustic features of vowels in two dialects of Kannada have been investigated which revealed that differences in duration of vowels were more in Mangalore than Dharwad dialects. F1 of /a/ and F2 of /i/ were greater for Mangaluru speakers than Dharwad speakers (Kapali, 2015). A study on variations in vowel space area in Mangaluru, Mysuru, Dharawad and Kalaburagi dialects of Kannada was performed. It was found that in Mangaluru dialect, F1 of /a/ and F2-F1 vowel space was the highest as opposed to other 3 dialects. Mangaluru dialect had the largest vowel space area (VSA) and Kalaburagi dialect had the smallest VSA. Also VSA was the second largest in the Mysuru dialect followed by the Dharawad dialect (Srinivasan, 2016). A similar study was done in dialects of Malayalam. It revealed that Ernakulam dialect had the largest VSA followed by Kozhikode and Trivandrum. VSA was the smallest in Thrissur dialect. The same study revealed that F_1 of /a/ and F_2 of /i/ was the highest in the Ernakulam dialect than other three dialects (Sarika, 2016). VSA was also studied in dialects of Telugu. It was reported that F1 of /i/ and /u/ were higher for Rayalseema speakers and F1 of /a/ was higher for Telangana speakers. Speakers of Coastal dialect had the highest VSA when compared to Telangana and Rayalseema dialects (Krishna and Rajashekar, 2012). Three Malayalam dialects (Trivandrum, Thrissur and Kannur dialect) were subjected to study for acoustic correlates. Samples were collected of speech and reading and PRAAT software was used for analysis. The speaking fundamental frequency, voice onset time (VOT) and temporal parameters such as syllable duration, vowel duration and word duration were analysed. The results revealed significant differences across the 3 dialects in few of the temporal parameters (M. Krishnan, N., Pringle, H., & T, Jayakumar, 2012). There is a possibility that such temporal differences could aid in the identification of a specific dialect. Hence, the current study attempts to provide an understanding of the role of temporal cues in identifying a dialect through a duration morphing paradigm.

Aim of the study

The aim of the study is to investigate the role of segmental / durational cues in the identification of a speaker's dialect.

Objectives

- To measure temporal aspects of the sentences in the speakers' dialects
- To measure the percentage of identification of the correct dialect in the duration morphed speech.

Chapter 2

REVIEW OF LITERATURE

Barkat, Ohala and Pellegrino (1999) carried out a study to investigate the efficacy of prosodic information in language identification and discrimination tasks in Arabic language. Speakers were selected from the two significant dialectal areas of the Arabian zones. Two adult male native speakers from Algeria and Morocco of the Western zone, and two adult male native speakers from Jordan and Syria of the Eastern zone were chosen for the study. A story telling task of each speaker was recorded. The stimuli consisted of twenty four passages which were to be presented to two populations of thirty eight adult listeners. One population had minimum or no knowlegde of Arabic dialects and the other population had prior knowlegde of the same. The fundamental frequency information was neutralised from the speech samples and presented to the listeners. The Arabic subjects identified correctly by fifty eight percentage accuracy and the non Arabic subjects identified correctly by forty nine percentage accuracy which was not significant. This proved that the Arabic listeners/speakers were aware of the the differences and were accurate in identifying eventhough the basic frequency information were removed.

Peters, Selting, Gilles and Auer; 2002 conducted a study on regional varieties of german languages which can be identified by intonational cues alone. Two experiments were performed. The first experiment involved the listeners to judge contours of Northern Standard German to contours of Hamburg urban vernacular. The second experiement involved listeners to judge contours of Berlin urban vernacular compared with both Low Alemannic German contour and Northern Standard German contour. The study revealed that listeners who were familiar to that dialect identified the contours. Non familiar listeners also identified the contours with some success rates. In the second part of the experiments, a carries phrase was added to the sentences. The first experiment involved a speaker who produced carrier phrases in Northern Standard German. The second experiment involoved two sets of carrier phrases. One set was from a Northern Standard German speaker and the other set was from a Berlin urban vernacular. When the carrier phrase was presented with the utterance spoken by the Berlin urban vernacular speaker, Berlin contours were recognised better. Yet, unirom effect was not found for the contour varities that were investigated. Whereas few studies have also shown poor identification rates.

Kehrein, Purschke and Lameli (2010) and Kehrein (2012) have revealed poor identification score for German dialects. A study was performed to investigate the role of prosody in synthesis and authentication of dialect. Yoon (2009) used Masan and Seoul dialect of the Korean language for the study. One male native speaker of Seoul dialect and two male native Masan dialect speaker were chosen. The Seoul utterances were modified by adding features of Masan dialect using prosody cloning technique. First, the sentences were manually segmented using PRAAT software. Eight Seoul sentences were modified by adding Masan segmental duration, Fo contour, intensity contour, duration & Fo contour , Duration & intensity contour, Fo contour & intensity contour; duration, Fo contour & intensity contour respectively. A three factor ANOVA analysis was performed. Main effects of the Fo contour and segemental durations were present but no interaction was found. The segmental duration, Fo contour and intensity transfer account for 67.3% of the total responses on a linear regresion analysis. The most responsible responses were provided by the absence or presence of the Fo contour transfer which was followed by the intensity transfer and duration transfer. Rao, K, Nandy, S, & G Koolagudi, S (2010) conducted a study using Auto-associative neural network (AANN) model to extract the spectral and prosodic characterisitcs from speech samples to distinguish between the dialects. Bengali, Chattisgharhi, Marathi, Telugu and General dialects of Hindi language were considered for the study. Speech sample is recorded from five female and five male speakers for each dialect. The participants were asked general questions thereby getting a spontaneous response. 5-10 minutes of speech samples were collected from each speaker. Using spectral, prosodic and combined spectral and prosodic features respectively, the average performance of the dialect identification system was 62%, 69% and 78%. The justification for the better performance of the combined system is attributed to the supplementary nature of the features.

Rao, K. S., & Koolagudi, S. G. (2011) used the Auto-associative neural network (AANN) models and Support Vector Machines (SVM) to extract the dialect specific information of Hindi language. AANN models extract the nonlinear relations which is specific to that dialect by using the distributions of feature vectors. SVMs execute dialect classification using discriminative characteristics which is present between the dialects. The prosodic and spectral characteristics were captured from speech samples for the identification of Chattisgharhi, Marathi, Bengali, Telugu & General dialects of Hindi. The investigations proved that prosodic features provide better dialect specific information than the spectral features.

Syiem, Deka, Ismail, Singh (2016) conducted a study using Gaussian Mixture model (GMM) to identify the two dialects of Khasi language spoken in Meghalaya.

Khynriem and Bhoi-Jirang dialects were chosen for the study. The speech corpus was collected from 8 female and 10 male speakers for each dialect. 10-20 minutes of spontaneous speech samples were collected from each speaker. Speech extraction methods and the Gaussian Mixture model were applied. Identification Rate (IDR) was used to determine the performance of the Dialect Identification System. The dialects were identified with 99% for the Khynriem dialect and 96% for the Bhoi- Jirang dialect. A similar study was done by Ismail and Singh using Gaussian Mixture Model (GMM) and Gaussian Mixture Model with Universal Background Model (GMM-UBM) to identify language and dialects. Goalparia and Kamrupi dialects of Assamese language were chosen for the study. Spontaneous speech samples were recorded from 21-60 years old speakers. Speakers were from both cities and villages. For Goalparia dialect, 3 hours from 27 speakers and for Kamrupi dialect, 4 hours and 22 minutes from 30 speakers were collected. The spectral Feature extraction was done using Mel Frequency Cepstral Coefficient (MFCC) and Fast Fourier Transform (FFT) was performed. The data gained after the feature extractions are installed into the GMM and GMM-USD for the training and testing phases. The results showed that both GMM and GMM-UBM identified the dialects successfully. The system identifies the dialects with 85.7% accuracy when GMM is used and identifies the dialects with 98.3% accuracy when GMM-USB is used.

Fuchs (2015) conducted a study to reveal if Indian English listeners could distinguish Indian English from British English and if yes which cues they relied upon. The mother tongue of the Indian speakers were Bengali, Malayalam, Tamil, Telugu and Hindi. Speech from 34 participants from which half of them were speakers of IndE and the other half were speakers of BrE were recorded using PRAAT software. 112 stimuli were presented to the listeners out of which 4 were original recordings and the others were resynthesised with the use of PRAAT's PSOLA algorithm. The stimuli were resynthesised using low pass filter, replacing the pitch contour and changing the duration of the vocalic and consonantal intervals. Participants had to listen to the stimulus one by one and say if the speaker was Indian or British. The segmental, intonation and rhythm contributed significant effects. However, it was found that segmental cues had the stronger influence on accent discrimination than intonation and rhythm.

Leemann and Siebenhaar (2008) conducted a study to investigate if prosody alone can act as a significant cue through which the regional origin of Swiss German speakers can be recognised. The Zurich, Bern, Valais and Eastern Switzerland dialects were chosen for the study. The experiment was carried out in two stages. In the first stage, 84 students of University of Zurich and 23 students of University of Berne were given 1 dialect sample of the 4 dialects. The spontaneous speech samples were collected from four male speakers from each dialect. The objective of this experiment was if the listners could recognise the four dialects without the sample being synthesised. For the second stage, the speech samples were modified. The fundamental frequencies were normalized to 140 Hz using PRAAT software and the pitch floor value and pitch ceiling value was changed. This data was filtered using a stop Hann band filter and smoothened at 500 Hz. So finally there were 4 unfiltered and 16 filtered speech samples. Two previously modified stimuli were included in the second stage. The timing and intonational characteristics of one phrase from a BE speaker was modified to that of a VS speaker and vice versa. The stimuli from both stages were presented randomly and were played only one time. The participants were instructed to indicate if what they heard was BE, VS, GR or ZH dialect on a questionnaire. Also, the specificity of their judgement were to be marked as ‘probably’ or ‘perhaps’. The results showed that GR and VS are recognised better by the ZH listeners in the GR case and by the BE listeners in the VS case. This is because of the considerable contact between BE speakers and VS speakers in Berne, & GR and ZH speakers in Zurich. 3 out of 4 dialects were identified above chance level. Furthermore, the BE and ZH listeners were poor in identifying their own dialects. Regarding the certainity of judgements, the listeners believed that the VS dialect was the easiest to be identified. The uncertain judgements for the other 2 dialects were also attributed because of the experimental condition- 5 seconds/judgement is a very small duration and also the subjects listened only once to each stimulus.

Vicenik and Sundara (2008) performed a set of five experiments investigated the adults’ ability to discriminate between American Southern Californian and Australia dialects by using different phonetic cues. Experiment 1 had full cue (rhythmic, intonation & segmental cues), Experiment 2 was low pass filtered speech (weakened segmental cues), Experiment 3 was flat intonation speech (rhythmic & segmental cues), Experiment 4 was crossed intonation (intonation; mismatched intonational & segmental cues), Experiment 5 had rhythmic and intonational cues only. 39 sentences were recorded from 8 Australian female speakers and 8 American Southern Californian female speakers. Sentences were modified using PRAAT. There were 10-14 native American English listeners. The sentences were played to them one by one. They were asked to identify as “American” or “Other”. The results revealed that the participants were able to use both prosodic and segmental information to differentiate their dialect from a foreign dialect. The accuracy of using segmental cues are higher than that of prosodic cues.

A study using the prosody morphing paradigm in Magrhrebian accented French was performed by Mareüil, Brahimi & Gendrot (2004). Speech samples were recorded from eighteen speakers of Kabyle, Arabic and French. They were approximately 25 years of age and equal number of males and females were present. The speakers resided in France for approximately 2 years. The speakers had to read a 400 word paragraph. The speech was recorded with a high quality microphone and digitized at 22.05 kHz, 16 bits, mono. The segmentation was done using PRAAT and prosody morphing done using PSOLA algorithm. The energy was normalised, and the suprasegmental and segmental features of different accents were morphed. The study showed that segmental factors aided in the perception of Magrhrebian accented French than prosody.

Leemann, Kolly & Nolan (2016) performed a dialect identification study in Bern (BE) and Valais (VS) Swiss German (SwG) dialects. This study focuses on the role of segmental and suprasegmentals cues especially the temporal cues. The speech sample was recorded from 6 BE speakers and 6 VS SwG speakers in the age range of 18-33. The sentences to be told were selected from the Bamford-Kowal-Bench (BKB) corpus. These were then manually marked in PRAAT to perform syllable based segmentation. The fundamental frequencies of both male and female were normalised. Then duration morphing was done. The syllable durations of sentence 1 of speaker BE1 were morphed onto the syllables of sentence 1 of VS1 speaker and vice versa. In a similar fashion the other 9 sentences were morphed. 30 Zurich German speaking listeners were chosen for the study. Listeners were made to get familiarized with the experiment. The stimuli were presented randomly for each listener. They had to decide if the sentence they heard was VS SwG or BE by clicking on the respective button on the laptop. Listeners also had to indicate their judgement accuracy on a scale ranging from 1= certain to 3= only guessed. The analysed data were represented in three conditions. The Effect of condition was significant that is in the unmorphed speech the listeners' sentivity was significantly higher. The unmorphed and the morphed condition did not show any difference in the listeners' confidence level. In the Between dialect difference condition, a higher % correct was shown in the unmorphed speech for both dialects. Confidence level scores decreased slightly in the morphed condition for VS SwG. In the Bias condition, there was wider spread of responses in the morphed condition indicating a more variation in the listeners preference towards either dialect. The study concluded that even though the syllable durations were exchanged between two dialects, listeners were able to recognise them. However there were variations.

Sereno, Lammers and Jongman (2016) performed a similar study in the perception of foreign accent in Korean accented English. Two female Seoul dialect Korean speakers and two native English speakers were chosen for the study to record speech samples. The 1st Korean speaker had been exposed to English for a total of 10 years and the 2nd Korean speaker has been exposed to English for 7 years. Degree of foreign accent were similar in both the speakers. The two English speakers were from the Midwestern American region and thus possessed the particular dialect of that region. 40 sentences from the Central Institute for the Deaf everyday sentences were recorded from the speakers. They read each sentence one by one. The recordings were made at 22050 Hz sampling rate. The mean Fundamental frequency (F0), maximum F0, minimum F0, F0 range and duration were measured for each of the sentences in PRAAT. Following this, each sentence was modified using the same software. The average amplitude was balanced to 70 dB. The same sentence spoken by two speakers: one provided intonation and the other segments; were paired before manipulating. The sample which provided the segmental information was lengthened or shortened in duration to match the intonation sample's duration. At the end, the F0 contour of segmental sample was morphed with the intonation sample's F0 contour. 40 native English speakers were included as listeners. The results revealed that segments had a significant effect on the accented speech perception than intonation.

Need for the study

Earlier studies in the Indian context have been carried out with emphasis on variation in acoustic characteristics in different dialects. The prospect of application of such studies has not been explored for dialect identification. Research in the past have not take into account particularly the role of durational cues in dialect identification. Most Indian studies have predominantly focused on using different models such as Auto Associative neural network models, Support vector Machine, Gaussian Mixture models to accurately identify different dialects using spectral or temporal cues.

These studies however has been automatized and not been carried out using an active listener's participation. They have not specified which acoustic cues majorly help in dialect identification, also not employed techniques such as prosody morphing paradigm. India being a multilingual country, dialect identification studies in the Indian context however have not been carried out across most languages. Malayalam is one such language in which extensive dialect identification research has not been carried out.

The above factors necessitate the present study. Malayalam is recognized in the Indian Constitution among the 15 official languages and is the mother tongue of twenty-two million people (Constitution of India, Schedule VIII. Out of a total population of 21,347,375 in Kerala; 20,496,778 people speak Malayalam. Subramaniam (1975) has appraised on the survey report which has reported about the twelve major dialect areas for Malayalam on the basis of the Ezhava caste dialect: South Travancore, Central Travancore, North Travancore, West Vempanad, Cochin, South Malabar, Central Malabar, North Malabar, South Eastern Palghat, North Western Palghat, Wayanad, Peak dialect. In the present study, Thrissur dialect which is spoken in Central Kerala and Kozhikode dialect which is spoken in the Northern part of Kerala will be chosen.

Chapter 3

METHOD

The present research was divided into two phases. The first phase aimed to compare duration of each syllable of same sentence across dialects. The second phase aimed to measure the percentage identification of the correct dialect in the duration morphed speech.

Phase I

The syllable duration of each sentence were compared across both the dialects.

Preparation of natural sentence stimulus

Dialects: Thrissur (T) and Kozhikode (K) dialects of Malayalam were selected for this study.

Materials: Sentences were selected as stimulus. These sentences had common words which were spoken in both dialects and varied only in suprasegmentals.

***Speakers*:** A female speaker aged from 18-30 of each dialect were chosen for the study. These speakers used the respective dialects on a daily basis. The speech samples were collected from their respective geographical region.

Inclusionary criteria:

- Speakers were native speakers of the respective dialects Thrissur and Kozhikode.
- They should not have exposure of any other dialect for more than 1 year.

- They should be above 18 years of age.

Exclusionary criteria:

- Participants with no history of hearing loss, speech, language, cognitive or neurological disorders.
- Also, participants with presence of structural or functional oro motor deficit were not included.
- Native speaker of that geographical region with exposure to other dialects were excluded from the study.

Procedure:

Speakers were seated in a comfortable seating position and the speech was recorded in a quiet room. Each sentence was given to the speakers prior to the recording for the purpose of familiarity and they were asked to read silently. After a time interval of 5 minutes, they were asked to speak out the sentences in their natural dialect. Olympus LS-100 recorder was used for recording. The Shure mic was used as an additional external mic. This omni directional microphone provides optimum speech intelligibility. The recording was done on sampling frequency of 44100 Hz and 16 bit rate. Each sentence was labeled.

Analysis

PRAAT was used to perform syllable segmentation. After this procedure, each syllable from the same sentences across dialects was measured for the syllable duration. For example: In the first sentence of dialect 1 (Thrissur), the duration of the syllable /vi/ was measured. Similarly, for dialect 2 (Kozhikode) duration of the same syllable of the same sentence was measured.

Phase II

A. Preparation of duration morphed stimulus sentences

In phase I, syllable duration of each sentences for the respective dialect was measured and tabulated. Using those syllable duration, the duration morphing was done between the dialects that is from Thrissur dialect, first sentence's first syllable duration value was morphed to Kozhikode dialect, first sentence's first syllable.

For example: Syllable duration of /vi/ of Thrissur sentence = 5 seconds

Syllable duration of /vi/ of Kozhikode sentence = 2 seconds

The 5 seconds of Thrissur syllable is changed to 2 seconds and vice versa.

B. Dialect Identification Task

Listeners

Inclusionary criteria:

- Listeners should have minimum 2 years of exposure to the above dialects mentioned.
- They should be above 18 years of age.

Exclusionary criteria:

- Participants with no history of hearing loss, speech, language, cognitive or neurological disorders.

Procedure:

30 listeners who were familiar with both the dialects were recruited for the study. Simple signal detection theory frame work was used for the listener

identification task. The listeners were instructed to decide each sentences are of which dialect. The stimulus was presented through professional headphones. The stimulus was presented three times. Out of the three, the average of the responses was taken for the analysis.

Statistical Analysis

The parameters obtained were tabulated and subjected to statistical analysis in Statistical Package for the Social Sciences (SPSS) software package (Version 21.0) and R2 statistical software package. McNemar's test was carried out to see the significant differences of number of correct responses between T-K to K-T with respect to each sentence. The percentage score is administered to the Shapiro-Wilk test for normality following which a parametric Paired t-test is carried out. Paired t-test is carried out to see the significant difference of total percentage correct response between T-K & K-T.

Chapter 4

RESULTS

The current study aimed at measuring the temporal aspects of the sentences in the speaker's dialect and quantifying the percentage of identification of the correct dialect in the duration morphed speech.

Temporal aspects of the sentence of the two dialects

The syllable duration in each sentence of both dialects were compared after syllable segmentation in PRAAT which is represented in the table 3.1. For example, the syllables of sentence 6 of Thrissur dialect was compared with the same syllables of sentence 6 of Kozhikode dialect (Figure 1) .

Figure 1: Comparison of syllable durations across both dialects

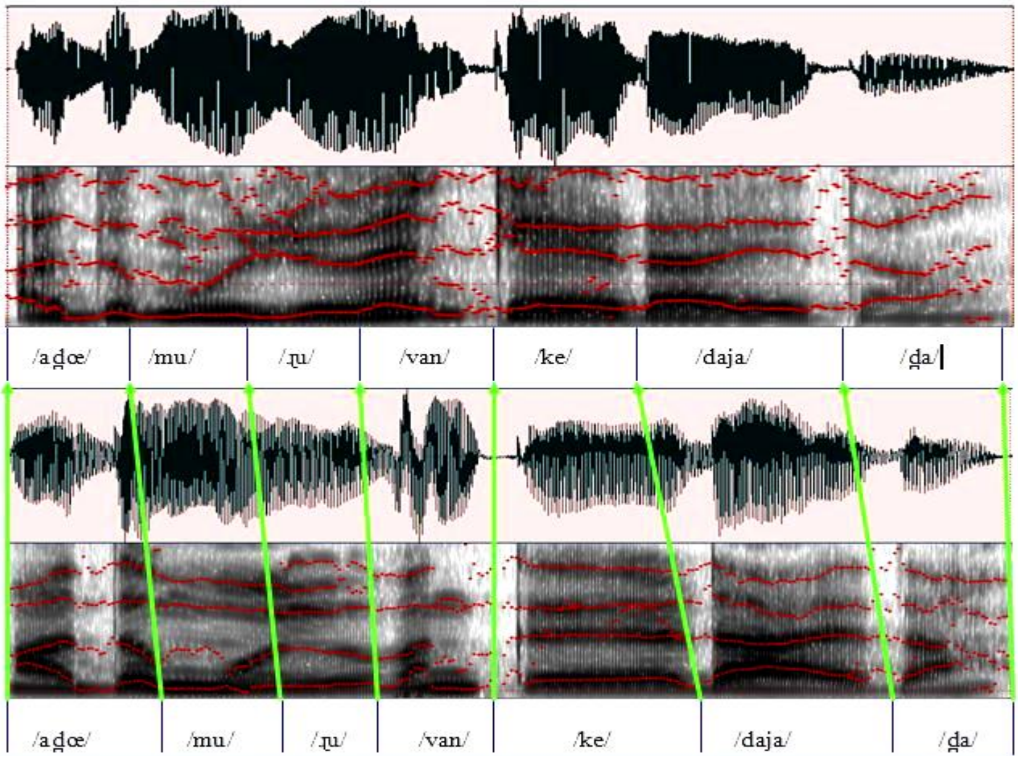

Table 1: Comparison of syllable duration of each sentence across dialects

Sl. No	Sentence	Thrissur (T) Duration (sec)	Kozhikode (K) Duration (sec)	Scaling factor T-K	Scaling factor K-T
1.	/mo/	0.162	0.165	1.018	0.981
	/l̥œ/	0.085	0.115	1.352	0.739
	/pɔɪ/	0.222	0.250	1.126	0.868
	/va/	0.317	0.09	0.533	1.869
2.	/ɪn̪n̪œ/	0.24	0.259	1.079	0.926
	/veɪ/	0.163	0.279	1.711	0.584
	/ɪla/	0.214	0.257	1.200	0.832
	/ɳœ/	0.161	0.081	0.503	1.987
3.	/na/	0.124	0.117	0.943	1.059
	/l̪l̪a/	0.174	0.246	1.413	0.707
	/t̪ɪ/	0.066	0.076	1.151	0.868
	/ra/	0.2	0.132	0.66	1.515
	/ka/	0.103	0.23	2.233	0.447
	/ɳa/	0.139	0.122	0.877	1.139
	/vɪ/	0.124	0.096	0.774	1.291
	/de/	0.121	0.079	0.652	1.531
4.	/na/	0.105	0.102	0.971	1.029
	/l̪l̪a/	0.259	0.216	0.833	1.199
	/bʱa/	0.068	0.153	2.25	0.444
	/ŋɪ/	0.153	0.243	1.158	0.629
	/ndœ/	0.211	0.229	1.085	0.921
	/to/	0.177	0.162	0.915	1.092
	/ka/	0.136	0.213	1.566	0.638
	/ɳan/	0.222	0.255	1.148	0.87
5.	/d̪eɪ/	0.167	0.221	1.323	0.755
	/vʌm/	0.8	0.084	0.8	1.25
	/el̪l̪/	0.249	0.236	0.947	1.055
	/am/	0.155	0.129	0.832	1.201
	/ka/	0.161	0.236	1.428	0.7
	/ɳun/	0.192	0.202	1.052	0.950
	/oɳdœ/	0.232	0.263	1.133	0.88
6.	/ad̪œ/	0.177	0.266	1.502	0.66
	/mu/	0.156	0.205	1.314	0.76
	/ɻu/	0.167	0.087	0.52	1.919
	/van/	0.201	0.192	0.96	1.041
	/ke/	0.20	0.266	1.33	0.751
	/daja/	0.306	0.262	0.856	1.167
	/d̪a/	0.242	0.189	0.780	1.280

7.	/tʃa/	0.194	0.255	1.314	0.760
	/ja/	0.108	0.115	1.064	0.939
	/kœ/	0.151	0.106	0.701	1.424
	/ku/	0.12	0.165	1.375	0.727
	/r̥a/	0.162	0.151	0.932	1.072
	/tʃœ/	0.079	0.11	1.392	0.718
	/ma/	0.188	0.209	1.111	0.899
	/d̪u/	0.048	0.051	1.062	0.941
	/ram/	0.17	0.175	1.029	0.971
	/ma/	0.107	0.118	1.102	0.906
	/d̪i/	0.208	0.274	1.317	0.759
	/ta/	0.276	0.231	0.836	1.194
8.	/ɲa/	0.123	0.113	0.918	1.088
	/ŋʌl̥/	0.175	0.212	1.211	0.825
	/ɪpo/	0.364	0.248	0.681	1.467
	/pu/	0.114	0.144	1.263	0.791
	/ra/	0.081	0.141	1.740	0.574
	/t̪œ/	0.13	0.162	1.246	0.802
	/poɪ/	0.171	0.196	1.146	0.872
	/va/	0.085	0.12	1.411	0.708
	/n̪i/	0.237	0.3	1.265	0.79
	/te/	0.085	0.193	2.27	0.44
	/ol̥u/	0.4	0.248	0.62	1.612

T-K= Syllable duration of **Thrissur dialect** morphed to **Kozhikode dialect**
K-T= Syllable duration of **Kozhikode dialect** morphed to **Thrissur dialect**
Scaling factor: Durational ratio between two dialect

In sentences 1, 2, 4, 5, 7, 8, the syllable durations of Thrissur dialect were lesser than that of the syllable durations of the Kozhikode dialect. In sentence 3 and 6, the syllable durations of Thrissur dialect were more than that of the syllable durations of the Kozhikode dialect. The overall results reveal that the rate of speech is faster in Thrissur dialect than Kozhikode dialect.

Percentage of identification of the correct dialect in the duration morphed speech.

Table 2: Number of listeners' responses across sentence

SENTENCES			K-T		Total
			K	**T**	
Sentence 1	**T-K**	**T**	8	3	**11**
		K	10	9	**19**
Total			**18**	**12**	
			K	T	
Sentence 2	**T-K**	**T**	17	5	**22**
		K	4	4	**8**
Total			**21**	**9**	
			K	T	
Sentence 3	**T-K**	**T**	15	9	**24**
		K	4	2	**6**
Total			**19**	**11**	
			K	T	
Sentence 4	**T-K**	**T**	14	8	**22**
		K	8	0	**8**
Total			**22**	**8**	
			K	T	
Sentence 5	**T-K**	**T**	15	8	**23**
		K	6	1	**7**
Total			**21**	**9**	
			K	T	
Sentence 6	**T-K**	**T**	13	8	**21**
		K	9	0	**9**
Total			**22**	**8**	
			K	T	
Sentence 7	**T-K**	**T**	18	5	**23**
		K	6	1	**7**
Total			**24**	**6**	
			K	T	
Sentence 8	**T-K**	**T**	8	7	**15**
		K	6	9	**15**
Total			**14**	**16**	

The results were carried out across sentences and across listeners. Table 2 shows the *listeners'* responses across sentences. The above table shows that in sentence 1 of T-K, 11 listeners' response was T and 19 listeners' response was K. Whereas K-T, 18 listeners' response was K and 12 listeners' response was T. Similarly, in sentence 2 of T-K, 22 listeners' response was T and 8 listeners' response was K. In K-T sentence, 21 listeners' response was K and 9 listeners' response was T. The other sentences can be interpreted in the same way. These responses are summarised in the table 3.

Table.3. Total responses of listeners for each sentence: Across sentences

Sentences	**T-K**			**K-T**		
	No.of listeners with response T	**% listeners with response T**	**No.of listeners with response K**	**No. of listeners with response K**	**% listeners with response K**	**No.of listeners with response T**
Sentence 1	11	36.7	19	18	60	12
Sentence 2	22	73.3	08	21	70	09
Sentence 3	24	80	06	19	63.3	11
Sentence 4	22	73.3	08	22	73.3	08
Sentence 5	23	76.7	07	21	70	09
Sentence 6	21	70	09	22	73.3	08
Sentence 7	23	76.7	07	24	80	06
Sentence 8	15	50	15	14	46.6	16

In all sentences except 1 and 8, for T-K, majority of listeners provided response as T. In sentence 8, equal number of listeners provided response as T and K. In all sentences except 8, for K-T, majority of listeners provided response as K. The response of sentence 8 was below 50%. These results clearly reveal that even after duration morphing of syllables of sentences, the majority of the listeners could identify the sentences of their own dialect.

Figure 2: Percentage of dialect identification: Across sentence

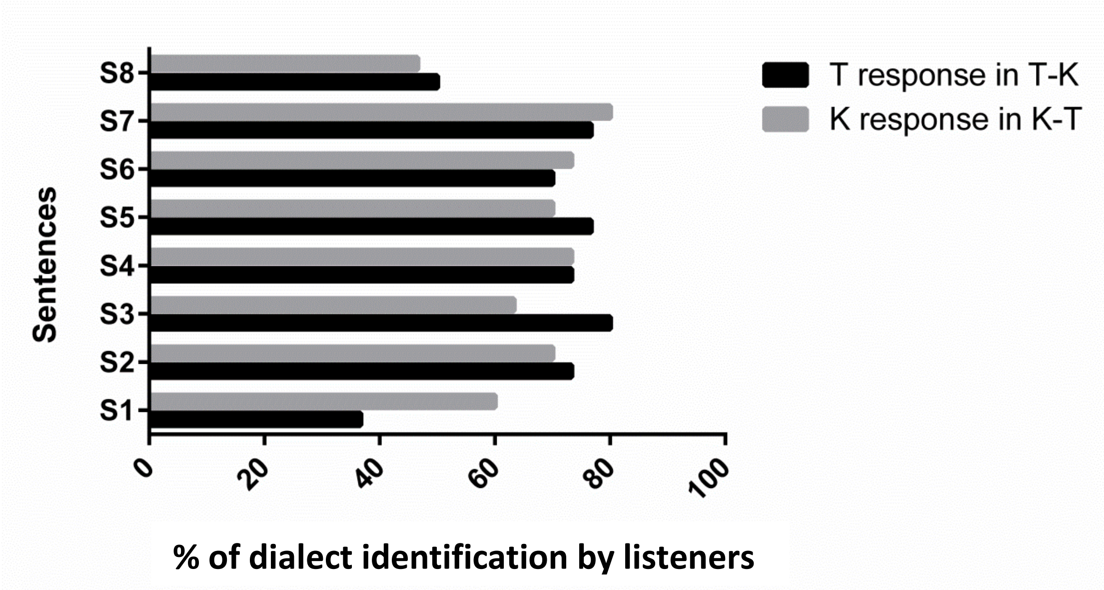

Table 4. Percentage of identification of the correct dialect in the duration morphed speech: Across listeners

Listener number	% of identification of T dialect in T-K morphed speech	% of identification of K dialect in K-T morphed speech
1	37.5	62.5
2	37.5	50
3	75	87.5
4	50	62.5
5	50	37.5
6	37.5	62.5
7	62.5	62.5
8	75	100
9	87.5	50
10	87.5	62.5
11	62.5	75
12	62.5	50
13	87.5	100
14	87.5	50
15	100	87.5
16	87.5	100

17	75	12.5
18	62.5	87.5
19	75	37.5
20	75	62.5
21	75	87.5
22	75	62.5
23	87.5	100
24	87.5	75
25	62.5	50
26	50	75
27	50	62.5
28	62.5	75
29	37.5	25
30	50	37.5

From the above table; In T-K condition, majority of the listeners except 1,2,6 and 29, identified the duration morphed sentences as T dialect. In K-T condition, majority of the listeners except 5,17,19,29 and 30, identified the duration morphed sentences as K dialect. These results reveal that majority of the listeners identified the sentences of their respective dialect (Thrissur or Kozhikode) in the duration morphed speech. The percentage scores were administered to the Shapiro-Wilk Test for normality. It follows normal distribution ($p>0.05$). Therefore a parametric paired t-test was carried out. The results revealed that there was no significant difference in the individuals' response in the duration morphed speech.

Table 5. t value and p-value of comparison of correct identification between T-K and K-T morphed sentences

	Mean	**N**	**SD**	**t**	**df**	**Sig. (2 tailed)**
T-K %	67.083	30	18.1241	0.507	29	0.616
K-T %	65.000	30	22.8375			

Chapter 4

DISCUSSION

The present study investigated the comparison of duration of each syllable of same sentence across dialects and aimed to measure the percentage identification of the correct dialect in the duration morphed speech. The results of the first objective revealed that the syllable durations of Thrissur dialect were lesser than those of Kozhikode dialect in 75% of the sentences. However mean effect of the syllable durations did not change the duration of sentences across dialects. This is displayed in table.1 in detail in the results.

Previous literature has shown such differences in temporal features in dialects. Jacewicz, Fox, and Salmons (2007) have reported vowel duration differences in three American English dialects (Inland North, South and Midlands). The results revealed that the South had the longest durations and the Inland North had the shortest. The midlands had an intermediate position. Miller, Mondini, Grosjean & Dommergues (2011) have shown differences in vowel duration in conjunction with the spectral information in Swiss French versus Parisian French. M. Krishnan, N., Pringle, H., & T, Jayakumar (2012) conducted a study to compare speaking fundamental frequency, voice onset time (VOT) and temporal parameters such as syllable duration, vowel duration and word duration in 3 dialects of Malayalam (Trivandrum, Thrissur and Kannur dialect). The results have shown significant differences across the 3 Malayalam dialects in few of the temporal parameters. Kapali (2015) has conducted a study in which it was revealed that differences in duration of vowels were more in Mangalore than Dharwad dialects of Kannada. Leemann, Kolly & Nolan, 2016 revealed that the Valais Swiss German speakers had a faster rate of speech than that of the Bern Swiss German dialect. Lohagun (2018) investigated on Nepali language. The results revealed that the standard variety Nepali speakers had longer vowel duration when compared to the other three variants (Sikkim and speakers from the Darjeeling and Alipurduar districts of West Bengal). This is contributed to both the effect of phonetic environment that is the effect of voicing and effect of region that is the front-back dimension of the vowel and height of the vowel on vowel duration.

Results of the second objective revealed that even in the duration morphed condition, there was no considerable change in the perception of dialect identification. It was also observed that the responses were variable from the same participant for the same sentences for the T-K morphed sentence versus the K-T morphed sentence as shown in table 2. This is in support with the previous literature (Leemann, Kolly & Nolan, 2016) which carried out a similar methodology by using PRAAT for syllable segmentation followed by duration morphing. This study involved Bern (BE) and Valais (VS) Swiss German (SwG) dialects. Even though the syllable durations were exchanged between the two dialects, listeners were able to identify the dialect. However dialect identification worsened significantly when durations were morphed. The variable confidence scores indicate that the response behaviour also becomes variable in the duration morphed speech which is related to mismatch in segmental and temporal information. Previous research by Rao. K. S. & Koolagudi. S. G (2011) have also given minimum emphasis on the role of durational cues in identifying a dialect. The prosodic and spectral characteristics were captured from speech samples for the identification of Chattisgharhi, Marathi, Bengali, Telugu & General dialects of Hindi. The investigations proved that the identification accuracy using duration is only 58% whereas using pitch, it is 69%. Whereas when all these features are combined the identification performance increases to 81%.

However in this study there were participants who identified the given stimulus as a different dialect in the duration morphed condition. 13% of the participants provided response as K in the T-K condition and 16% of the participants provided response as T in the K-T condition. Similar results were found in other studies. Vicenik and Sundara, 2008 revealed that the accuracy of using segmental cues are higher than that of prosodic cues. The study reports that for discrimination, prosodic cues can be used but there is low overall accuracy. Mareüil, Brahimi & Gendrot (2004) also revealed that segmental factors have a more prominent role than prosody. This is attributed to the fact that when speakers have a strong accent, the segmental factors override other prosodic features.

Also, in this study the frequency was not normalised which can again contribute to the poor change in the perception of dialect identification. The percentage of change in perception might have increased if other acoustic features like frequency, intonation, stress pattern etc were also neutralised. This is in support with the previous literature. Mareüil, Brahimi & Gendrot (2004) revealed that the increase in identification rates were also attributed to normalisation of the energy. In the study by Leemann, Kolly & Nolan (2016), the identification rates had increased when the fundamental frequency of both the males and females were normalised. This also prevented listeners from identifying speakers based on the frequency.

Implications

- This study would aid in increasing knowledge if segmental information has a role in speaker dialect identification.
- Such detailed dialectal studies will enable to develop an efficient Automatic Speaker Recognition System (ASR) for dialect identification.

Limitations

- In the synthesized speech output, slight distortions were present due to poor quality of the instrument. Using a better instrument would have provided with more clearer samples which further would have increased the quality of perception.
- Another limitation of the study is that before presenting it to the listeners, only the amplitude was normalised and normalisation of the pitch was not done. Normalisation of pitch variation (intonation) would have increased the percentage of change in identification of dialect.
- Also the language proficiency of the listeners were not formally investigated.

Chapter 5

SUMMARY AND CONCLUSIONS

Studies on dialectal differences have emerged from around the 1900s. From then several developments have occurred. Differences were identified in many languages in the aspects of linguistic features and acoustic features. However, there were hardly any studies which explored specific cues and role of that in dialect identification. Malayalam being the mother tongue of twenty-two million people, such measures was not investigated. Therefore Thrissur and Kozhikode dialects were selected for the study. The two phases of the study involved comparison of duration of each syllable of same sentence across dialects and followed by measurement of percentage identification of the correct dialect in the duration morphed speech.

In Phase I, same sentences from both dialects were recorded by speakers of that dialect. The recorded sentences were then labeled in PRAAT and segmented in terms of syllables. Then, each syllable from the same sentences across dialects were compared for the syllable duration. In Phase II, syllables of each sentence of dialect 1 (T) were duration morphed to the same syllable of same sentence of dialect 2 (K). 30 listeners familiar with both dialects were asked to identify dialect of each duration morphed sentence. Statistical analysis was performed using McNemar's test and Paired t-test.

The result of the first objective revealed that six out of 8 sentences of Thrissur dialect had reduced syllable duration when compared to that of Kozhikode dialect.

The second objective revealed that though the syllables of sentences were duration morphed, majority of the listeners could identify the sentences of their own dialect. Statistical analysis revealed no significant difference in the individuals’ response in the duration morphed speech. This can be attributed to the fact that in the present study, the frequency and other acoustic feature like intonation and rhythm was not neutralised. Frequency information mostly contributed to the identification of that dialect through identification of the prosody features of the speaker voice. The intonation information contributed to the better identification of that specific dialect. This study showed that there is role of durational cues in dialect identification. However it was not statistically significant. The identification percentage may be affected by other pitch variation cues in the sentences

REFERENCES

- Sapir, Edward. (1921). *Language, an introduction to the study of speech.* New York: Harcout, Brace and company.

- Kumari, Preethi. R. (2009). *Dialect diversity and linguistic change in a Kerala village.* Ph.d dissertation, University of Kerala.

- Robert, J. (2006). *Studying Dialect.* Macmillan Education, UK.

- Petyt, K. M. (1980). *The study of dialect: An introduction to dialectology*. UK: Westview Press.

- Bin Li. (1951). Identifying /l/ and /n/ in English by multi-dialectal speakers of Mandarin. *The Journal of the Acoustical Society of America, 143*(3).

- Matthew J. Makashay. (1999). Dialectal differences in diphthong perception. *Journal of the Acoustical Society of America, 106*(4), 2273.

- Cynthia, G., Clopper, David B., Pisoni., & Kenneth de Jong. (2005). Acoustic characteristics of the vowel systems of six regional varieties of American English. *The Journal of the Acoustical Society of America,118*(3), 1661–1676.

- Diwakar Mishra & Kalika Bali. (August 2011). *A Comparative Phonological Study of The Dialects Of Hindi.* Paper presented at the International Congress of Phonetic Sciences, Honkong.

- Bipasha Patgiri. (November 2011). *Stress in the Nalbaria Dialect of Assamese.* Paper presented at the International seminar on prosodic interfaces.

- Subramaniam, V. I. (Ed.). (1993). *Dravidian Encyclopedia* (pp. *487).* International School of Dravidian Lingistics. Thiruvananthapuram.

- Jayaraj, T. (2017). The Realisation of the Phonological Variation of /ɻ/ in Malabar Mappila dialect of Malayalam. *Jadavpur Journal Of Languages and Linguistics,01*(02), 039-048.

- Jacewicz, E., Fox, R. A., & Salmons, J. (2007). Vowel Duration In Three American English Dialects. *American Speech, 82*(4), 367-385.

- Chládková, K., Escudero, P., & Boersma, P. (2011). Context-specific acoustic differences between Peruvian and Iberian Spanish vowels. *The Journal of the Acoustical Society of America, 130*(1), 416-428.

- Abramson, A. S., & Ren, N. (1990). Distinctive vowel length: Duration vs. spectrum in Thai. *Journal of Phonetics*, *18,* 79–92.

- Miller, Joanne L., et al. "Dialect Effects in Speech Perception: The Role of Vowel Duration in Parisian French and Swiss French." *Language and Speech*, *54*(4), 2011, 467–485.

- Lohagun, Nishant. "Comparing Formants and Vowel Duration in Standard and Non-Standard Varieties of Nepali." *Language in India, 18*(4), Apr. 2018, 225–238.

- Kapali, N. (2015). *Acoustic characteristics of vowel of dialects in Kannada*. Unpublished Master's Dissertation, AIISH, Mysore.

- Srinivasan, Sahana. (2016). *Vowel space area across different regional dialects of Kannada.* Unpublished Master's Dissertation, AIISH, Mysore.

- Sarika, Rini. (2016). *Vowel space area across different regional dialects of Malayalam.* Unpublished Master's Dissertation, AIISH, Mysore.

- Krishna, Y., & Rajashekhar, B. (2012). Vowel Space Areas across Gender, Age and Dialects in Telugu. *Language in India, 12*(1), 357-369.

- Leemann, A., Siebenhaar, B., Germanistik, I., & Leipzig, U. (2008). Perception of Dialectal Prosody, Department of Linguistics , University of Berne , Switzerland. *Perception*, 524–527.

- M. Krishnan, N., Pringle, H., & T, Jayakumar. (2012). Acoustical Difference of Dialectal Variation in Malayalam Languages. In *National Seminar on Speech Synthesis and its Applications*.

- Barkat, M., Ohala, J., & Pellegrino, F. (1999). Prosody as a distinctive feature for the discrimination of arabic dialects. *6th European Conference on Speech Communication and Technology (EUROSPEECH'99)*.

- Peters, J., Gilles, P., Auer, P & Selting. M. (2002). Identification of Regional Varieties by Intonational Cues. An Experimental Study on Hamburg and Berlin German. *LANGUAGE AND SPEECH*, *45*(2), 115-139.

- Kehrein, R., Lameli, A., Purschke, C. (2010). *"Stimuluseffekte und Sprachraumkonzepte,"* in: C. Anders, M. Hundt, and A. Lasch (Eds.), "Perceptual dialectology". Neue Wege der Dialektologie. Berlin/New York: de Gruyter, 351–384.

- Kehrein, R. (2012). "Wen man nicht alles für einen Sachsen halt?! Oder: Zur Aktivierung von Sprachraumkonzepten durch Vorleseaussprache," in: K. Jakob, R. Hünecke (Eds.), Die obersächsische Sprachlandschaft in Geschichte und Gegenwart. Heidelberg: Winter, 223- 263.

- Rao, K. S., Koolagudi, S. G., & Nandy, S. (2010). Identification of Hindi dialects using speech, The 14th World Multi-Conference on Systemics, Cybernetics and Informatics.

- Yoon, K. (2009). The role of prosody in dialect synthesis and authentication. *Phonetics and Speech Sciences, 1*(1), 25-31.

- Rao, K., Nandy, S., & G Koolagudi, S. (2010). Identification of Hindi Dialects using Speech. In *14th World Multiconference on Systemics, Cybernetics and Informatics (WMSCI)*.

- Rao, K. S., & Koolagudi, S. G. (2011). Identication of Hindi Dialects and Emotions using Spectral and Prosodic features of Speech. *Systemics, cybernetics and informatics, 9*(4).

- Syiem, A., Deka, G. K., Ismail, T., & Singh, L. J. (2016). Khasi Dialects Identification Based on Gaussian Mixture Model. *International Journal of Engineering Science and Computing,6*(4).

- Ismail, T., & Singh, L. J. (2017). Dialect Identification of Assamese Language using Spectral Features. *Indian Journal of Science and Technology,10*(20).

- Fuchs, R. (2015). “You’re not from around here, are you? – A dialect discrimination experiment with speakers of British and Indian English,” in: E. Delais-Roussarie, M. Avanzi, S. Herment (Eds.), Prosody and language in contact: L2 acquisition, attrition and languages in multilingual situations. Frankfurt: Springer, 121-146.

- Leemann, Adrian & Siebenhaar, Beat. (2008). *Perception of Dialectal Prosody*. Paper presented at the 9th Annual conference of the International Speech Communication Association, Brisbane.

- Vicenik, C., & Sundara, M. (2008). The role of segmental and intonational cues in dialect discrimination. *The Journal of the Acoustical Society of America*, *123*(5).

- Leemann, A., Kolly, M.-J., & Nolan, F. (2016). Identifying a speaker's regional origin: the role of temporal information. Paper presented at the Conference: *Speech Prosody*, Boston, USA.

- Boula de Mareüil, P., Brahimi, B., & Gendrot, C. (2004). Role of segmental and suprasegmental cues in the perception of Maghrebian-accented French. In *INTERSPEECH 2004 - ICSLP, 8th International Conference on Spoken Language Processing*. Jeju Island, Korea.

- SERENO, J., LAMMERS, L., & JONGMAN, A. (2016). The relative contribution of segments and intonation to the perception of foreign-accented speech. *Applied Psycholinguistics*, *37*, 303-322.

APPENDIX

1. മോള് പോയി വാ.
/moḷœ/ /pɔɪ/ /va/

2. ഇന്ന് വെയിലാണ്.
/Inn̪œ/ /veɪlaɳœ/

3. നല്ല തിരക്കാണവിടെ.
/nal̪l̪al/ /t̪Irakaɳavɪde/

4. നല്ല ഭംഗിയുണ്ടല്ലോ കാണാൻ.
/nal̪l̪a/ /bʰaŋIndœto/ /kaɳan/

5. ദൈവം എല്ലാം കാണുന്നുണ്ട്.
/d̪eIvam/ /el̪l̪am/ /kaɳunonḍœ/

6. അത് മുഴുവൻ കേടായതാ.
/ad̪œ/ /muɻuvan/ /kedajad̪a/

7. ചായക്ക് കുറച്ച് മധുരം മതിട്ടോ
/tʃajakœ/ /kuɾatʃœ/ /mad̪uram/ /mad̪ita/

8. ഞങ്ങടൾ പുറത്ത് പോയി വന്നിട്ട് ഉള്ളു
/ɲaŋaḷɪpo/ /puɾat̪œ/ /poɪ/ /van̪ite/ /oḷu/