PERCEPTION OF COARTICULATORY CUES BY INDIVIDUALS WITH NORMAL HEARING AND COCHLEAR HEARING LOSS IN THE PRESENCE OF NOISE

Shreyank P Swamy

13AUD022

This Dissertation is submitted as part fulfillment
for the Degree of Master of Science in Audiology
University of Mysore, Mysore
MAY-2015

CERTIFICATE

This is to certify that this dissertation entitled "**Perception of coarticulatory cues by individuals with normal hearing and cochlear hearing loss in the presence of noise"** is a bonafide work in part fulfillment for the degree of Master of Sciences (Audiology) of the student (Registration No. 13AUD022). 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.

Mysore

May, 2015

Dr. S. R. Savithri
Director
All India Institute of Speech and Hearing
Manasagangothri, Mysore- 570 006

CERTIFICATE

This is to certify that this dissertation entitled "**Perception of coarticulatory cues by individuals with normal hearing and cochlear hearing loss in the presence of noise"** is a bonafide work in part fulfillment for the degree of Master of Sciences (Audiology) of the student (Registration No. 13AUD022). This has been carried out under my guidance and has not been submitted earlier to any other University for the award of any other Diploma or Degree.

Mysore

May, 2015

Dr. Sandeep M.
Guide
Reader in Audiology
Department of Audiology
All India Institute of Speech and Hearing
Manasagangothri, Mysore- 570 006

DECLARATION

This is to certify that this dissertation entitled "**Perception of coarticulatory cues by individuals with normal hearing and cochlear hearing loss in the presence of noise"** is the result of my own study under the guidance of Dr. Sandeep M., Reader in Audiology, Department of Audiology, All India Institute of Speech and Hearing, Mysore and has not been submitted earlier to any other University for the award of any other Diploma or Degree.

Mysore

May, 2015

Registration No.: 13AUD022

Dedicated to

Appa, Amma

&

my remarkable guide

Sandeep sir ☺

Acknowledgments

As the famous saying goes "a journey of thousand miles starts with a single step". 6 years back I stepped into this esteemed institute and today I can feel that I have not only developed as a professional but as a person in whole. I owe this success to few important people in my life and now I would like to take this opportunity to acknowledge all of them.

First and foremost I would like to thank God almighty for his blessings, for showing me the path and guiding me always. My parents are the greatest blessing in my life. I am grateful to you appa and amma for being amazing parents. Appa- you have been my mentor, my inspiration and have always motivated me in every walk of life. Thank you so much for everything. Amma- you are an epitome of love, care, and selflessness. You have accepted all my temper tantrums with a smile and helped me through all the difficult times. Thank you amma for being the world's best amma ☺. I would also like to thank my dear grandparents, brother, relatives and all well-wishers for all their support.

My remarkable and awe-inspiring guide – Sandeep sir. You have not only taught me the "Essentials of Audiology" but also the "Essentials of Life". You have been my guiding light from the time I stepped into this place. I admire your splendid persona, your spirit and demeanor. Thank you sir for being so patient with me, answering my silliest questions and maintaining your tranquility and composure even in the most difficult times. Thank you so much sir for being there always ☺

Priya ma'am you have always welcomed us with a smile. A sincere thanks for the warm hospitality. Late night coffee and all time ready snacks were truly the energy boosters. Thank you mam ☺. A special thanks to the little angel Anvita- you have always brought a smile on my face.

A sincere note of thanks to Asha ma'am, Animesh sir, and Ajith sir for being brilliant teachers who always instigated my mind. Your knowledge and wisdom has no bounds.

Sharath sir- without you this study would have never been possible. Your magical voice will always be acknowledged. You have always been there to help and teach. Thank you sir ☺

I would also like to express my gratitude for Chandini ma'am and Devi ma'am who helped me to rate and validate the stimulus. Also I would like to thank all the participants of the study for their valuable time and patiently sitting throughout the testing.

I would like to thank my teachers- Yeshoda ma'am, Goswami sir, Jayashree ma'am, Gopishankar sir and Ganapati sir for being amazing teachers and for all their support throughout.

I would also like to express my gratitute to Gnanvel sir, Antony sir, Megha mam, Arun sir, Prashanth sir and Vikas sir for all their timely help and support.

Nike anna- you have been my search engine I would say. I have always irritated you with my endless doubts and you have always had answers to them and helped me throughout. Thankyou anna ☺. Deepashree akka- thank you so much for being there

at the right moment and helping me. Srikar anna- thankyou for all your support. A sincere thanks to Kavitha ma'am.

And here comes a big thanks to the library staff- I don't think any of us can ever finish any of our assignments without your excellent management and ever-readiness to help search books and related materials.

Mohana akka- you have always been an elder sister to me and taken care of me. Thank you for your love and compassion. I cherish all the moments we spent together. Prasanna Hegde (my college brother)- you have always been a source of inspiration for me. You are my partner in adventure. Thank you anna ☺

Divya – You have always been there for me. Thank you so much for every thing you have done for me. Thank you for tolerating me. I'm very lucky to have YOU in MY LIFE. I won't say much, because I know you don't need words to understand me.

Spruha (mankad) – you have always played dual role of a friend and a sister. I am blessed to have you. All the fun times, silly fights, calling names, and pulling leg, everything has added innumerable lovely moments to my life. Thank you so much ☺

My amazing swim team: Raksha, Pooja, Hemanth, Manjuganesh and Monoj- early morning chills, diving into the freezing pool, and bonus sprints for being late, we have all made it fun together. Thank you guys. My childhood buddies – Nithin and Sharath, thank you for making school always exciting and fun. I would also like to thank all my teachers, classmates, juniors and seniors at Vidhyashala and Gurukula.

My seniors who have been more like a friend- Suppi (Lilliput), Indu (Korangi), Lakshmi akka, Sandhya (gube), Ansu, Neethu, Rajan and Sabarish, thank you all for all the good times. I will always cherish these days.

Preeta- my study partner or rather I would say my teacher. Studying with you has always been fun. You have always added flavors to the boring subjects ☺

Aditi, Mani, Pankaja, Rekha, Sonam, Vinodhini and Yashu- thankyou guys for all the support, wonderful time and unlimited fun.

My li'l naughty sisters: Varsha, Veepee and Jasi thank you so much for all the crazy talks and laughs we shared. Janani (Gundu) and Deepthi thank you both too.

A special note of thanks to NSS which has helped me grow as a person.

This acknowledgment will be incomplete without thanking my 'Expendables' team who made sports always a driving force.

Last but not the least I would like to thank all my B.Sc and M.Sc classmates.

Thank you one and all ☺

Abstract

The effect of background noise, direction of coarticulation and cochlear hearing loss, on the perception of stop consonants based on coarticulatory cues was studied. The stimuli were unvoiced stop consonants /p, t, k/ in vowel context /a, i, u/. Twenty normal hearing and 12 SNHL participants were tested at 40 dB SL in quiet and 10 dB SNR in both CV and VC syllables. All the CV and VC syllables were truncated into 3 tokens each i.e., tokens with full transition + vowel (VT), half transition + vowel (VHT) and only vowel (V). Results revealed that, full transition with vowels served as the most useful carryover coarticulatory cue. Spread of anticipatory coarticulation cues was greater than carryover cues and was significant for consonant perception. Further, addition of noise reduced anticipatory coarticulatory perception more than carryover coarticulatory perception. In SNHL group, perception based on coarticulatory cue was poorer than normal hearing group.

Table of Contents

List of Tables

List of Figures

Chapter 1

INTRODUCTION

The influence of articulators of one phonological segment on the adjacent phonological segment/s is known as coarticulation. According to Kuhnert and Nolan (1999), a phonological segment is not realized identically in all environments, but often apparently varies to become more like an adjacent or nearby segment secondary to coarticulation.

Ohde and Sharf (1981) defined coarticulation in three different aspects that is physiologically, acoustically and perceptually. '*Physiologically*, it refers to the integration of neural commands to the speech musculature, timing and movement patterns of articulators and aerodynamic forces. This results in the spreading of features from one sound to another. *Acoustically***,** it refers to acoustical modifications in consonants and vowels by the contextual phonemes. *Perceptually***,** it refers to the difference in the place, manner and voicing features of phonemes perceived by the individual, due to the influence of adjacent phonemes.

The coarticulation can occur with the influence of consonants on vowels or vice versa. The perception of coarticulation is determined by the acoustic cues such as formant transitions in identifying the vowel or consonant from consonant-to-vowel (CV) and vowel-to-consonant (VC) combinations. Many researches using synthesized speech have indicated formant transition as the major acoustic cue for the identification of consonants (Cooper, 1952; Ohman, 1965). Cooper (1952) reported that for place of articulation of the consonants, second formant (F2) transition between the consonant and vowel is shown to be the important cue, whereas Wang (1959) found that both transition and release cues play role in the perception of final plosives. Dubno et al. (1987) reported short-duration spectral cues such as voicing duration is also important for identifying the consonants when paired with vowels. Voicing duration serves as an important cue for hearing loss subjects to identify the consonants, as the spectral onset cue gets distorted in them. It was seen that as the voicing duration increases, consonant recognition scores in /a/ and /u/ combination in hearing loss subjects approached normal hearing listeners, but the scores for vowel /i/ combination remained low for hearing impaired individuals.

Coarticulation can be characterized in terms of several aspects such as its direction and degree. Direction of spread of coarticulatory cues (Daniloff & Moll, 1968) is of two kinds, anticipatory effect and carryover effect. In anticipatory effect, the sound segment influences a preceding sound. While in carryover effect, the sound segment influences the following sound. The anticipatory effect is because of phonemic preplanning and is reported to prevail for longer time, whereas the carryover effect is because of phoneme utterance and its effect is larger than anticipatory effect.

Various methods have been used in studying speech perception. Ohde and Sharf (1977) recorded CV and VC syllables studied the order effect of acoustical segments in identifying the stops and vowels. The spectrograms of the original CV and VC utterances were measured in millimeter. Obtained values was converted into time and then into recording tape length. The obtained CV and VC syllables were separated into *aperiodic, aperiodic+ vocalic transition, vocalic transition and vocalic transition+ vowel segments* through tape deletion method. It was observed that consonant identification scores were greater for CV syllables compared to VC syllables in *vocalic transition+ vowel segments*, but no significant difference were seen in other 3 types of tape deletion. But for vowel identification, scores was greater in *aperiodic segments* for CV syllables compared to VC syllables and no significant difference was seen in remaining tape deletion methods. *Bhuvaneshwari (1993)* recorded four voiceless consonants in medial position of CVCV words to study the development of perception of coarticulation. Five synthetic stimuli were generated for each word using DWSSLC software were; original word (-vbtv), original word with the burst removed (-Vtv), stimuli from the beginning of the initial consonant till the burst of the key stop (-Vcb), stimuli from the beginning of the initial consonant till the transition of the key stop consonant by removing the burst using cut and splice technique (-Vt), and the stimuli from the beginning of the initial consonant till the transition of the key stop consonant. The steady portion of the end vowel was removed using the cut and splice technique (-Vcbt). It was observed that responses were declined in the order of –vbtv, -vtv, -vcbt, -vt and –vcb, which suggests vowel transition plays a major role in perception of consonants. Lehiste and Shockey (1972) studied perception of coarticulatory effects in English VCV syllables. The VCV syllables were cut into two parts by placing the cut in the voiceless plosive gap using splicing technique which results in VC and CV sequences. Both VC and CV sequences were not perceived in the deleted segments. They concluded that coarticulatory cues may hinder the perception of non-final allophones when artificially placed in final position.

1.1. Justification for the Study

Considering that consonants contribute primarily to the intelligibility of the speech, if one has to understand speech, it is important to perceive the consonantal cues. However due to their low amplitude nature, consonantal cues are likely to get missed in the adverse listening conditions such as in the presence of noise. In such instances coarticulatory cues of consonants present in the vowels tend to contribute for the consonantal speech identification. However, such facilitation by coarticulatory cues of consonants on vowels shall be dependent on the signal-to-noise ratio (SNR) and shall have its own limitations beyond a particular SNR. Although the review of the literature reveals that coarticulatory cues are helpful in the speech identification it is not quite clear as to whatis its perceptual utility in the presence of noie signal to noise ratio, these coarticulatory play facilitative function in the consonantal identification. The effect of noise on coarticulation perception may also depend on the direction of coarticulation, coarticulatory cue per se (For example, part or complete transition) and the presence or absence of cochlear hearing loss. However such an attempt to probe into the usefulness of coarticulatory cues in noise has not been done in the past. The findings from such an experiment will be of both academic and clinical interest.

1.2. Aim of the Study

To document the effect of background noise on perceptual utility of coarticulatory cues.

1.3. Objectives of the Study

The objectives of the study are:

1. To compare the perception of consonants based on co-articulatory cues, between quiet and 10dBSNR
2. To compare the consonant identification scores between anticipatory and carryover coarticulatory conditions, in quiet and 10dBSNR.

3. To compare the consonant identification scores between normal hearing and cochlear hearing loss individuals, in quiet and 10dBSNR.

Chapter 2

REVIEW OF LITERATURE

Speech perception was an enigma for a long period till researchers in the field started exploring and discovering various aspects involved in the processing of speech sounds. Coarticulation being one of the essential components of speech perception and processing, has attracted the attention of many researchers who have abundantly contributed to the literature regarding coarticulation cues in speech perception.

Various studies have been conducted to study the coarticulation perception with different methods. One of the most frequently used methods in the earlier days were pattern playback technique (Liberman, Delattre, & Cooper, 1958) and tape deletion method (Ohde & Sharf, 1977) to study the speech perception. Many authors used this kind of synthesized method to study the coarticulatory effect of vowels and consonants, where it was easy to synthesize and study the cues which are responsible for the perception of consonants from vowels or vice versa.

2.1. Coarticulatory Perception in Normal Hearing Individuals

Many researches using synthesized speech have indicated formant transition as the major acoustic cue for the identification of consonants (Cooper, 1952; Ohman, 1965). Cooper (1952) reported that for place of articulation, second formant (F2) transition between the consonant and vowel is shown to be the important cue, whereas Wang (1959) found that both transition and release cues play a role in the perception stop consonants within final positions.

The perception of stop consonants depends on the frequency position of the burst and the F2 transition. Based on the relative placement of burst frequency with reference to F2 frequency, it is perceived as /k/ or /p/. When the burst frequency is low, if the following vowel /i/ or /u/ it would be perceived as /p/ and with vowel /a/ it is perceived as /k/. Therefore, although we can categorize stop consonants it depends on the vowel contexts (Cooper, 1953).

Schatz (1954) reported that unvoiced stops /p, t, k/ depends more on coarticulatory cues than on primary cues. She compared the perception of stops in different vowel contexts in case of synthetic speech experiment using playback machine and actual speech experiment using cut and splice method. In both the experiments the results revealed similar findings as, for the burst frequency >3000 Hz, it was perceived as /t̲/ regardless of vowel contexts. When the burst frequency was <3000 Hz, vowel context along with burst frequency was required to perceive as /p/, /t/, or /k/. Halle et al. (1957) reported that on truncation of final consonants identification of /p, t, k, b, d, g/ depends on the vowel contexts.

Sharf and Hemeyer (1971) reported formant transition as the important cue in identifying consonant in CV and VC. They studied the importance of the CV and VC formant transitions in the identification of place of articulation. They tape recorded labials /p, b, f, v/, alveolar /t̲, d, s, z/, and palatal /k, g, ʃ, ʒ / with vowel /ə/ combination to study the perception of CV and VC transitions in voiced stops, unvoiced stops, voiced fricatives and unvoiced fricatives. The results revealed that more correct identification of consonants was in VC than CV formant transitions. This finding was supported by Brady, House and Stevens (1961). In CV formant transitions fricatives are better identified than stops, whereas vice-versa happened in VC transitions.

Sharf and Hemeyer (1971) listed three models for the identification of consonants better in VC formant transition than CV: transition direction, transition dependence, and transition sufficiency model. In the *transition direction model*, it depends on the direction of change in frequency and rate of change in frequency. The direction of change in frequency in the CV transition reaches or leans upon the vowel steady state which is the terminal frequency of the transition. Whereas in VC transition it reaches the consonant. This terminal frequency provides more information on the vowel in CV and about final consonants in VC (Brady, House & Stevens, 1961). In the *transition dependency model*, the information provided by the vowel transition for consonant identification is based on the information provided by the burst frequency of the stops. The burst frequencies of the initial stops are more important cue for perception than the burst frequency of the final stops. House et al. (1965) support this model by reporting; at various signal-to-noise ratios initial stops are better perceived than final stops. In American English pronunciation, the final stops bursts are unreleased. Because there is insufficient pronunciation of the final consonant, the listener depends more on VC transition than CV transition. In the *transition sufficiency model,* VC formant transition provides better information on place of articulation than CV formant transition. In CV transition, part of the transition is represented in the burst duration of the consonant which leads to difficulty of stop perception when the burst frequency is truncated in the CV than VC. This leads to difficulty in making decision for CV transition than VC transition. This may be due to phonetic assimilation were the effect of forward coarticulation is greater than backward coarticulation.

Sharf and Beiter (1974) investigated identification of consonants from vocalic transition+ vowel segments by presenting in forward and backward direction to test the above models. The results revealed identification of consonants from vocalic transition+ vowel segments were greater for VC than CV. But this was exception in case of fricatives in both forward and backward conditions, which also decreases in errors for CV condition and increase in errors for VC condition in backward direction. This implicates the possibilities of masking effect of vowels on consonants cues. When the stimulus is played in backward direction, there may be change in masking phenomenon in the opposite direction. These results supported the transition sufficiency model, whereas the backward condition supports the transition dependency model and transition direction model.

Based on the findings of the Sharf and Hemeyer (1972), and Sharf and Beiter (1974), Massaro (1973) reported that the difference in perception of CV and VC syllables may be due to the temporal order relationship between the steady-state of the vowel and the vocalic transition in the syllables. The identification of consonant is better for VC vocalic transition due to the following silent period in the syllable compared to CV vocalic transition, which is followed by the steady-state of the vowel resulting in masking effect on the burst amplitude of the consonant. This highlights importance of temporal order in speech perception.

The finding by LaRiviere, Winitz, and Herriman, (1975) and Wang (1959) on vocalic transition contradicts the transition dependency model. LaRiviere et al. (1975) reported no significant difference between aperiodic only and aperiodic+ vocalic transition stimuli in both CV and VC syllables. They that stated aperiodic segment is a strong cue and vocalic transition provides insufficient or no cue for perception of unvoiced consonants. But the influence of vowel on consonant provides a strong coarticulatory cue for the perception of aspiration. Wang (1959) reported that, with the release cues, identification of the final consonants is better than only with the transition cues. He also reported that duration of the preceding vowels, formant transitions, duration of the voicing or silence following transition and the burst frequency are important for the perception of final consonant in a monosyllabic context.

Lehiste and Shockey (1972) studied perception of coarticulatory effects in English VCV syllables using the consonants /p, t, k/ and vowels /i, a, u, ae/. The VCV syllables were cut into two parts by placing the cut in the voiceless plosive gap using splicing technique which results in VC and CV sequences. Results showed difficulty in consonant identification when the VCV stimuli were spliced into VC context than CV. They concluded that coarticulatory cues may hinder the perception of non-final allophones when artificially placed in final position. Also they hypothesized saying ‘coarticulatory effect reduces when the phoneme/ speech sound is placed in an environment where they do not occur naturally’.

Ohde and Sharf (1977) recorded CV and VC syllables using the stops /p, t, k, b, d, g/ and vowels /a, i, u/ to study the order effect of acoustical segments in identifying the stops and vowels. The spectrograms of the original CV and VC utterances were measured in millimeter and values were converted into time and then into recording tape length. The obtained CV and VC syllables were separated into aperiodic, aperiodic+ vocalic transition, vocalic transition and vocalic transition+ vowel segments through tape deletion method. Results revealed that consonant identification scores were greater for VC syllables compared to CV syllables in vocalic transition+ vowel segments, but no significant difference were seen in other three tape deletions.

Pols and Schouten (1978) investigated the identification of the deleted initial and final consonant with and without 300 ms noise burst stimuli preceding (CV) or following (VC) the stimuli. The consonants used in the study were /p, t, k, b, d/+ /i, u, a/+ /t/ in the initial position and /t/+ /i, u, a/+ /p, t, k/ in the final position. Results revealed identification of consonant with VC transition is better than CV transition. Because, when a burst is truncated from the CV or VC stimuli it results in click like sensation. When the click is in the beginning of the transition, the scores of the CV transition reduces compared to VC transition as the click occurs at the end results in burst like perception. When noise burst were introduced near the click in case of VC transition, identification of consonants did not vary much. But in case of CV transition, identification of consonants improved with noise. This may be due to masking effect of the noise burst over the click like sensation. Also they report response bias towards the identification of /p/ for the deleted consonants in both CV and VC conditions even in the presence of noise. This may be due to perception of the click, as burst of consonant /p/.

Bhuvaneshwari (1993) recorded four voiceless consonants in medial position of CVCV words to study the development of perception of coarticulation using /p, t, k/ consonants. Five synthetic stimuli were generated for each word using DWSSLC software; original word (-vbtv), original word with the burst removed (-Vtv), stimuli from the beginning of the initial consonant till the burst of the key stop (-Vcb), stimuli from the beginning of the initial consonant till the transition of the key stop consonant by removing the burst using cut and splice technique (-Vt), and the stimuli from the beginning of the initial consonant till the transition of the key stop consonant. The steady portion of the end vowel was removed using the cut and splice technique (-Vcbt). It was observed that responses were declined in the order of –vbtv, -vtv, -vcbt, -vt and –vcb, which suggests that the vowel transition plays a major role in perception of consonants.

Modarresi, Sussman, Lindblom, and Burlingame (2004) investigated the acoustic properties of the VCV utterances in stops to study the anticipatory and carry-over coarticulatory effects in open and closed syllables. They found greater anticipatory effect in labials than carry-over effect, whereas it was vice versa in case of alveolars and velars. In closed syllables, carry-over effect was higher than anticipatory effect in the following order velar> alveolars> labials. In open syllables, labials and velars showed higher anticipatory effect than carry-over effect, but in alveolars, carry-over effect remained higher than anticipatory effect. They reported, that stop place as a function of anticipatory or carry-over effect is based on the vowel combination. In case of alveolars, carry-over effect was seen in /a, i, u/, but in case of vowel /e/ context anticipatory effect was higher than carry-over effect.

2.2. Speech Perception in Adverse Listening Condition in Normal hearing listeners

Gordan-Salant and Wightman (1983) investigated the perception of stops in presence of different types of noise in normal hearing group. Results revealed in the presence of noise the consonant identification were affected in the following sequence: velars > alveolars> bilabials.

Wang and Bilger (1973) investigated the effect on consonant identification in CV and VC condition in the presence of BBN across different SNRs (-10 to 15 dB

SNR) in normal hearing individuals. Results revealed performance was improved with the increase in SNR. Further, the performance was better for CV condition compared to VC condition.

Leibold and Buss (2013) studied the consonant identification in the presence of speech shaped noise and 2 talker babble noise at -10 and 0 dB SNR. This study was investigated on normal hearing children and adults. Results revealed in both children and adult group, consonant identification were affected in the following sequence: place > manner or voicing in 2 talker babble and manner or place or voicing in speech shaped noise at -10.

Finitzo-Hieber and Tillman (1978) studied the effect of 8 talker babble on speech perception in different SNRs (0, +6, +12 dB SNR & quiet). Results revealed that, as the SNR increased from 0 to 12 dB, performance become better. Performance at +12 dB SNR was close to quiet condition.

2.3. Acoustic Cues for Consonant Perception in Individuals with Hearing Impaired

The speech perception errors of the individuals with hearing impairment can be explained through description of their phonemic composition and by examining the acoustic cue structure of the misperceived phonemes. The important reason which explains speech perception deficits is the audibility of acoustic cues of phonemes. The audibility depends on the degree and configuration of hearing loss. The audibility is relatively lesser for consonants compared to vowels and hence they suffer from forward and backward masking from high intensity vowels, when audibility is reduced.

For consonants there is a higher load on temporal processing due to shorter durations and dependence on transitions. Along with this, finer frequency resolution is important for consonant perception.

In conductive hearing loss, speech perception is mainly due to inaudible phoneme acoustic cues. In individuals with SNHL and other neural pathologies, have even more degraded performance due to inaudible cues, reduced frequency selectivity (Moore, & Glasberg, 1986; Moore, 2001; Kluk & Moore, 2005), loss of compressive nonlinearity (Moore, 1998), abnormal loudness growth (intensity coding) (Moore, 1998), poor temporal processing (Viemeister, 1979; Hopkins & Moore, 2009) and pattern recognition (central) processes. Studies report that, in SNHL place of articulation (POA) is affected, as second formant frequency and the direction of formant transition is the major cue for POA. Along with POA, perception of manner and voicing is also affected due to poor temporal processing which leads to difficulty in encoding rapid formant transitions, VOT, and duration of preceding vowel (Boothroyd, 1984).

Boothryod (1984) investigated the degree of hearing loss at which the consonant identification scores fell to 50%. He found that with increase in the degree of hearing loss the cues for consonant identification were affected in the following sequence: POA (75 dB HL) > voicing cues (85 dB HL) > manner cues (90 dB HL).

Van Tasell, Hagen, Koblas, and Penner (1982) investigated the differences across normal hearing individuals and individuals with mild hearing loss in terms of utilizing the spectral cues of place of articulation. Their stimuli consisted of 14 full length natural syllables and 14 synthesized syllables on a continuum of /ba-da-ga/. Keeping F1, F4, and F5 constant, they varied the F2 and F3 on the continuum. Results revealed a good correlation in the accuracy scores for two conditions i.e. full length natural syllables and full length synthetic syllables. Accuracy scores for full length synthetic syllables were reported to be 100 % for normal hearing and 83- 100 % (mean= 89 %) for individuals with SNHL. For the short stimuli conditions, /ba-da-ga/ boundary was displaced to the right for individuals with hearing loss, /ga/ being the most common response. Thus, they concluded that there were no significant differences found between the two groups in terms of using spectral cues to identify place of articulation. This could be because of the lesser degree of hearing loss considered.

Turner and Rob (1987) reported that in individuals with hearing impairment, an increase in intensity lead to better audibility. However, their speech perception scores were below the normal level. Different patterns of responses observed were: 100% audibility and near normal recognition scores, 100% audibility and below normal recognition scores, and less than 100% audibility and near normal recognition scores. They concluded that audibility of individual stop consonants influences recognition performance in individuals with SNHL.

2.4. Coarticulatory Perception in Individuals with Hearing Impaired

As discussed above there are several cues required for perception of stops consonants. Processing of speech sounds by human auditory system involves extraction of spectral and temporal cues which aid in perception and discrimination of different speech sounds. However, any damage or impairment in the auditory system which affects the hearing abilities of an individual would also lead to impaired perception of speech sounds. Coarticulation is an important phenomenon which influences the perception of speech sounds in different contexts. Thus, it becomes essential to understand how these coarticulatory cues influence speech sound perception in individuals with hearing impairment.

In individuals with SNHL, POA, voicing and MOA are affected as the degree of hearing loss increases (Boothroyd, 1984). As the degree of hearing loss increased individuals with SNHL showed reduced perception of consonant voicing due to which phoneme confusion occurs (Pickett et al., 1972). Revoile et al. (1982) investigated the voicing perception for final stops in 2 groups (normal hearing & HI- moderate-severe). They considered the stops /p, t, k, b, d, g/ with different cue modifications in several conditions. Findings revealed that during exchange of vowel ongoing transitions between the cognate pairs, vowel transition was an important cue for perception of voicing in final stops. In case of unvoiced stops, burst deletion resulted in decreased voicing perception for both the groups. Later in 1987, Revoile et al. extended their study to investigate voicing perception in initial stop perception. They reported VOT was a very important cue for voicing perception in both hearing impaired and normal hearing listeners.

Dubno, Dirks, and Schaefer (1987) reported that, short-duration spectral cues such as voicing duration is also important for identifying the consonants when paired with vowels. Voicing duration serves as an important cue for hearing loss subjects to identify the consonants, as the spectral onset cue gets distorted in them. It was seen that as the voicing duration increases, consonant recognition scores in /a/ and /u/ combination in hearing loss subjects approached normal hearing listeners but, the scores for vowel /i/ combination remained low for hearing impaired individuals.

Hedrick et al. (1995) reported that individuals with hearing impairment depend more on relative amplitude information when compared to formant transition to decide the place of articulation. The task given to the participants was to judge the syllables presented as /pa/ or /ta/, when presented at three relative amplitude i.e. -10, 0, and +10 dB. Results of experiment 1 (without noise) revealed that normal hearing individuals used formant transitions as a cue for judging place of articulation unlike individuals with SNHL who depended more on relative amplitude of the burst than formant transition cues. It was also observed that individuals in both the groups had fewer labial responses with increase in presentation level. The labials were found to be replaced with alveolar responses in both the groups, more so for individuals with SNHL. Increased frequency of alveolar response among individuals with hearing loss was explained as a function of recruitment phenomenon. This recruitment phenomenon in the impaired ear increases the relative amplitude of burst than the vowel in high frequencies and leads to more alveolar responses as alveolars are known to have more high frequency energy than labials (Odhe & Steves, 1983). In the second experiment, normal hearing individuals were tested in the presence of noise to elevate their threshold and were observed if their performance equates with the performance of individuals with SNHL. It was found that both the groups performed differently. This difference can be explained by drawing support from the findings of Dubno and Schaefer (1992). They stated that masking noise might not be the ideal simulation as it creates extra stimulation which is not present in the impaired ear. Based on their findings, authors concluded that the differences between the two groups cannot be explained based on the differences in their audibility.

Hedrick and Jesteadt (1996) carried out another investigation and reported that in contrast to manipulation of relative amplitude only, significant difference was present in the performance of normal hearing and SNHL individuals when both relative amplitude and formant transition cues were manipulated. They justified their findings by drawing support from perceptual weighting hypothesis which states that, the two groups assign different weight to the two acoustic cues. They report that due to lack of audibility of formant transitions, different weights were assigned by the individuals with SNHL. Another explanation given was disrupted formant transition coding in individuals with SNHL. They also investigated the influence of vowel duration on labial perception between the two groups. Finding revealed that unlike normal hearing individuals, increase in vowel duration did not increase the labial responses in individuals with HI. This indicated that the dynamic processing of relative amplitude in SNHL is not as time dependent as in case of normal hearing individuals. This is attributed to reduced temporal integration observed in individuals with SNHL.

2.5. Speech Perception in Adverse Listening Condition in Individuals with Hearing Impaired

Dubno et al. (1982) investigated the consonant confusion in noise for individuals with flat, gradual sloping, and steeply sloping hearing loss using nonsense syllables. Results revealed that the group with gradual sloping hearing loss performed better followed by flat and steeply sloping hearing loss groups. Whereas, with respect to consonant perception higher errors for place of consonant compared to manner and voicing were noticed. In consonants, better scores were obtained for voiced consonants than unvoiced consonants, with no difference between CV and VC conditions. In case of place of consonants all the three groups showed greater scores for velars compared to alveolars and labials.

Helfer and Huntley (1991) compared the consonant recognition in older adults with minimal HL to moderate high frequency HL and young adults with normal hearing in the presence of cafeteria noise using nonsense syllables. Between the two groups in manner of articulation, plosives and fricatives were most affected and nasals were least affected. In place of articulation, bilabials and dentals were most affected in both the groups with better scores in young adults. In case of older adults even alveolars and palatal errors were present which was absent in young adults.

Beattie, Barr, and Roup, (1997) investigated the effect of SNRs (+5 to +15 dB SNR) on speech perception in the presence of 12 speaker multi talker babble in normal hearing and mild to moderate hearing loss individuals. Results revealed improvement in speech identification scores with the increase in SNRs in both the groups. Keith (1972) and Stelmachowicz et al. (1985) also reported the similar findings in flat and high frequency hearing loss groups.

Gardan-Salant (1987) studied the influence of hearing loss and aging on consonant recognition in the presence of speech babble (+6 dB SNR). The findings revealed that performance of normal hearing elderly individuals was better than individuals with hearing impairment. However, the pattern of consonant perception could not be related to their audiometric configuration.

Chapter 3

METHOD

The aim of the present study was to document the effect of background noise on perceptual utility of coarticulatory cues. Factorial research design was used to compare the clinical group i.e. individuals with sensorineural hearing loss (N= 10) and control group i.e. individuals with normal hearing (N= 20). It was hypothesized that there is no significant difference in the perception of coarticulatory cues by individuals with normal hearing and cochlear hearing loss in the presence of noise.

3.1. Participants

A total of 30 native speakers of Kannada who know to read and write in Kannada participated in the study. All the participants were in the age range of 20 to 50 years and were geographically from in and around Mysuru. The test battery used for screening and selection of participants in both the groups consisted of pure tone audiometry to determine the type and degree of hearing loss, immittance evaluation to rule out presence of any middle ear dysfunction, DPOAEs to determine outer hair cell functioning and QUICKSIN to rule out the presence of any auditory processing deficits or retro cochlear pathology. Based on their hearing sensitivity, the participants were devided into two groups; normal hearing group and the sensorineural hearing loss (SNHL) group.

***Group I*:** It included twenty normal hearing individuals with the hearing thresholds within 15 dB HL at octave frequencies between 250 Hz and 8 kHz. All the participants had type-A tympanogram (tympanic peak pressure ranging between -100 daPa to +60 daPa, with a static admittance of 0.5 to 1.75 ml) and acoustic reflexes were present at 0.5k Hz, 1k Hz, and 2k Hz in both ipsilateral and contralateral conditions. The participants were also screened for cochlear dysfunction by ensuring the presence of DPOAEs with SNR greater than +6 dB at three consecutive frequencies.

***Group II*:** It consisted of 12 individuals with acquired SNHL since past 3 to 12 months. The type of hearing loss i.e. cochlear or sensorineural was determined based on the air bone gap (ABG < 10dB). All the participants had mild degree of hearing loss (average of pure tone thresholds ranging between 26 to 40 dB HL) with flat or gradual sloping audiogram pattern. All the participants had type-A tympanogram (as defined earlier) with acoustic reflexes present or absent depending on the degree of hearing loss. None of the participants in the group II were rehabilitated for hearing loss.

The participants of group I were age and gender matched to the participants in group II. Informed consent was taken from all the participants and they were all assured of maintenance of confidentiality. The details of the participants of the study are given in Appendix 1.

3.2. Test Stimuli

The present study considered stop consonants as they are characterized by maximum number of acoustic cues when compared to all other consonant types. This redundancy of cues helps in better perception compared to other speech sounds. The test stimuli included consonant-vowel (CV) and vowel-consonant (VC) syllables. These syllables consisted of one of the three unvoiced stop consonants /p/, /t/, and /k/ paired with one of the three vowels /a/, /i/, and /u/. This resulted in a total of 9 CV and 9 VC syllables (Appendix 2). This was done in order to assess anticipatory and carryover coarticulation in the three vowel contexts.

These syllables were uttered by an adult male who was a native speaker of Kannada. He was instructed to produce each of the syllables 4 times at normal loudness and with a clear articulation. A condenser microphone kept at a distance of 6 inches away from the speaker's mouth and was connected to MOTU Microbook II as an external module for better quality of recording. The recording was performed using Adobe audition (version 3) in a personal laptop computer in a sound treated room. The stimuli were digitized and each syllable was recorded 4 times with a sampling frequency of 22 kHz and 16 bits quantization and was saved as wave files.

The 4 samples of the recorded syllables were given to 5 individuals to rate the CV and VC syllables for their intelligibility and clarity of sound. These individuals were Audiologists/Speech Language Pathologists with an experience of more than 5 years in the field. The ratings were given in a 5 point scale (1-very bad, 2- bad, 3- average, 4- good, 5-very good). The syllables with rating of 4 or 5 were selected as final set of syllables for synthesizing. The syllables /pa/, /pu/, /ap/, /ip/, /up/, /ik/, and /uk/ were rated less than 3 in terms of intelligibility. The listeners reported perceived air puffs for VC syllables with /p/ and poor burst intelligibility for VC syllables with /k/. Therefore, these syllables were re-recorded and goodness rating was repeated till the syllables with ratings more than 4 were obtianed. Later, the entire selected stimuli were normalized to RMS value using Adobe audition to maintain equal intensities across the syllables.

3.2.1. Technique of Splicing the Speech Stimuli

The normalized syllables (CV and VC) were edited using Praat software to systematically remove the consonantal cues for assessing the coarticulatory perception. Each of the 18 syllables were spliced to create the following 3 stimulus tokens using Praat software: vowel + full transition, vowel + half transition and only vowel component (consonant was excluded). After this stage, the total numbers of stimuli were 72 (18x4): 1 set of complete syllables + 3 sets of modified tokens.

In the present study the formant transition segment of CV or VC were determined by drawing a horizontal line over the steady state of F2 formant and vertical line at the point where F2 formant deviates from the horizontal line as shown in Figure 3.1. The duration of the transition was obtained from the point of vertical line to the point of onset (CV) or offset (VC) of the periodicity. Using this, various segments of the required speech sounds which are included in the present study i.e. full transition+ Vowel (steady state) (VT), half transition+ vowel (steady state) (VHT) and only vowel (V) were obtained. This method was adapted from Odhe and Sharf (1977).

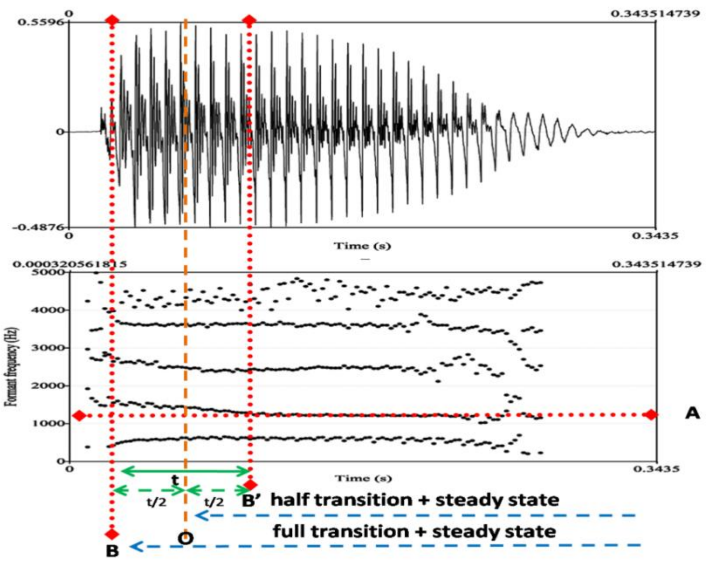


Figure 3.1: *W*aveform in top panel and formants in bottom panel of syllable /ta/. In the bottom panel A horizontal line represents the steady state of F_2 formant, B' vertical line represents the offset of the formant transition, B vertical line represents the onset of the formant transition or the periodicity. The segment between B-B' denoted as 't' represent the full transition segment. O vertical line represents the half transition which was obtained by dividing t/2.

3.2.2. Procedure of superimposing noise on syllables

In order to assess the role of coarticulatory cues of CV and VC syllables in the presence of noise, the complete syllable and the synthesized truncated were mixed with speech shaped noise in terms of RMS signal-to-noise ratio at +10 dB SNR using MATLAB program (Gnanateja & Pavan, 2012). Considering the control 'quiet' condition (72 tokens) and the +10 dB SNR (72 tokens), each individual was presented a total of 144 stimuli.

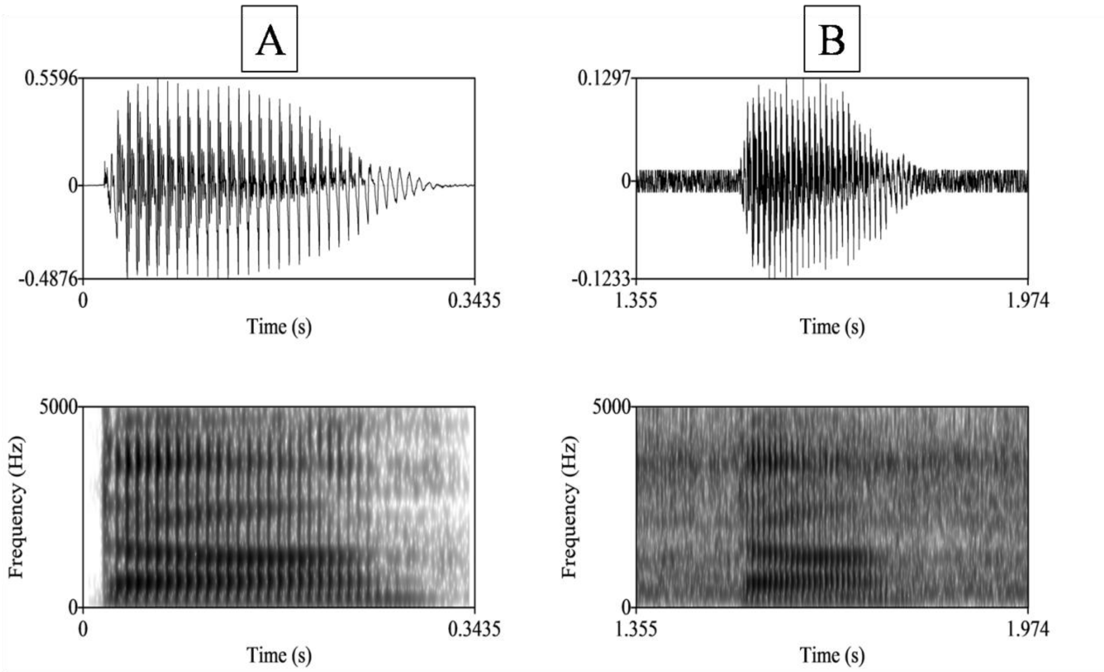


Figure 3.2: Shows the waveform in top panels and spectrogram in bottom panels. A shows consonant /ta/ in Quiet condition and B shows consonant /ta/ superimposed with noise (+10 dB SNR).

3.2.3. Design of Paradigm Software for Stimulus Presentation and Response Recording

The paradigm software was custom designed to present the stimulus tokens in a random order and to record the respective individual responses. Each type of stimulus token (complete syllable, full transition + vowel, half transition + vowel, and only vowel) was presented as separate list in each SNR (Quiet & 10 dB SNR) and in each type of syllable (CV & VC). Using this software the audio files were linked with the numerical keys of the laptop computer. The numeric key 1 was linked to /pa/, 2 was linked to /ta/, and 3 was linked to /ka/. Further the mouse pointer on the screen was also linked to the displays in the graphic user interface (prepared using Paradigm software) which was /pa/, /ta/, and /ka/. The software was programmed to score the response as 1 or 0 depending upon its correctness. It also recorded the response consonant (/pa/ or /ta/ or /ka/ depending on the numeric key pressed by the individual) in each stimulus presentation. For example, if an participant identified /pa/ as /ka/, the response was scored 0, and the response consonant was stored as /ka/.

The software was also programmed to control the inter-stimulus interval. The inter-stimulus interval was kept as 3 seconds. However, the next stimulus was presented only after the response key was pressed for the previous stimulus presented.

3.3. Test Administration

All the participants were given the following instruction for the Quiet condition: “Please listen carefully to the sounds presented. Sounds will be presented only to your right ear. As you listen to the sound, you need to identify which among /p/, /t/, or /k/ is presented and either press the respective key or use the mouse pointer to click on the syllable in GUI. While testing at 10 dB SNR an additional instruction was given to ignore the noise and identify the consonant.

The entire listening test was presented at 40 dB SL **(**Keni et al., 1979) in an audiometric room with noise levels within permissible limits according to ANSI S3.1-1999 (R2003). The stimuli were played using Paradigm player in a personal laptop computer through calibrated head phones (Seinhessier HD 180). The participants were instructed to listen carefully to the stimulus and identify the perceived consonant. The 3 consonants (/p/, /t/, /k/) were displayed on the computer screen and the participants had to choose one of them (closed set identification) either by clicking on the intended consonant using the mouse or by pressing the specified keys.

Each speech token was presented thrice in each list randomly and also the order of the lists of speech tokens was randomized across the participants. Each list had 27 speech tokens and there were 16 lists. The testing was carried out in multiple sessions with breaks as required. Table 3.1 represents the number of lists presented to each individual.

3.4. Response Analysis

Each response was scored 0 for incorrect and 1 for correct response. This was stored in a Microsoft Excel sheet. The total number of correct response for each consonant, in each speech token, in CV and VC syllables, in Quiet and 10 dB SNR was calculated. The group data was then analyzed to identify the coarticulatory cues for each consonant. The data was analyzed statistically using appropriate tests.

Table 3.1: *Lists of test stimuli used in the present study*

Context	SNR	List No.	Lists
Consonant Vowel	Quiet	1	Complete syllable
		2	Full transition + Vowel
		3	Half transition + Vowel
		4	Only Vowel
	10 dB SNR	5	Complete syllable
		6	Full transition + Vowel
		7	Half transition + Vowel
		8	Only Vowel
Vowel Consonant	Quiet	9	Complete syllable
		10	Vowel + Full transition
		11	Vowel + Half transition
		12	Only Vowel
	10 dB SNR	13	Complete syllable
		14	Vowel + Full transition
		15	Vowel + Half transition
		16	Only Vowel

Chapter 4

RESULTS

The aim of the present study was to document the effect of background noise on perceptual utility of coarticulatory cues. In the present study, group served as an independent variable, while SNR, stimulus condition, CV v/s VC combination served as within subject factors. The dependent variable was the consonant identification scores. The effect of within-subject factors and the group on the dependent variable was studied using Mann-Whitney U test, Friedman test and Wilcoxon test.

Mann-Whitney U test was used to compare the identification scores between the two groups, Friedman test was used to compare the identification scores across the 4 stimulus tokens (CV, VT, VHT & V), and Wilcoxon test was used to compare between Quiet and 10 dB SNR. Wilcoxon test was also used for pair-wise comparisons in instances where Friedman test showed a significant main effect. The results thus obtained are reported under the following headings:

1) Identification of coarticulatory cues
2) Effect of noise on the perception of coarticulatory cues
3) Effect of carry-over versus anticipatory coarticulatory cues on the perception of stop consonants
4) Effect of group on coarticulatory perception in Quiet and 10 dB SNR

4.1. Identification of Coarticulation Cues

The primary aim of the present study was to test the effect of noise and hearing loss on the perception of coarticulatory cues. Therefore, the first task in this process was to identify the cues that provide coarticulatory information for perception of consonants. For this, only the data of the normal group in the Quiet condition was used. Results of the same are reported separately for CV and VC syllables.

4.1.1. Results of CV Syllables

To identify the cues that provide adequate coarticulatory information, a 1 out of 3 criteria (33.33%) was fixed. That is, if the mean scores of the consonant identification was more than 1, only then the cue was considered to be providing adequate coarticulatory information for perception. On the other hand, if the mean score in a particular speech sound was 1 or less than 1, cue was not considered to be providing adequate coarticulatory information. Figure 4.1 shows the mean and SD of consonant identification scores of normal hearing participants in the four stimulus tokens (CV, VT, VHT, & V), in Quiet condition in CV syllables. Figure 4.1 shows that the mean scores as a general trend, decreased as the cues were truncated. This was true for most of the syllables in all the three places of articulation.

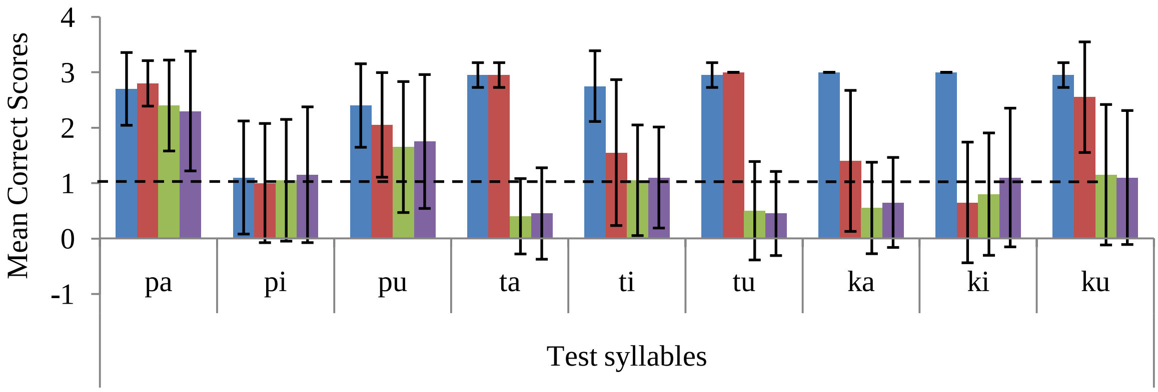


Figure 4.1: Mean and SD of consonant identification scores of normal hearing participants in the four stimulus tokens (CV, VT, VHT, & V), in Quiet condition in CV syllables. Mean scores below the dashed line were considered to be not conveying the coarticulatory information.
Note: Maximum individual score was 3.

The horizontal dashed line demarcates between the cues that provide adequate coarticulatory information and the cues that do not provide adequate coarticulatory information. In CV syllables based on the one out of three criteria, several VHT and V tokens (except in /pa/ and /pu/) did not qualify to be the coarticulatory cues for the consonant identification. However, VT in all the syllables except for /pi/ and /ki/ qualified to be coarticulatory cues for the perception of consonants.

4.1.2. Results of VC Syllables

Similar to CV syllables, coarticulatory cues for the perception of consonant identification were identified in the VC syllables. Figure 4.2 shows the Mean and SD of consonant identification scores of normal hearing participants in the four stimulus tokens (VC, VT, VHT, & V), in Quiet condition in VC syllables.

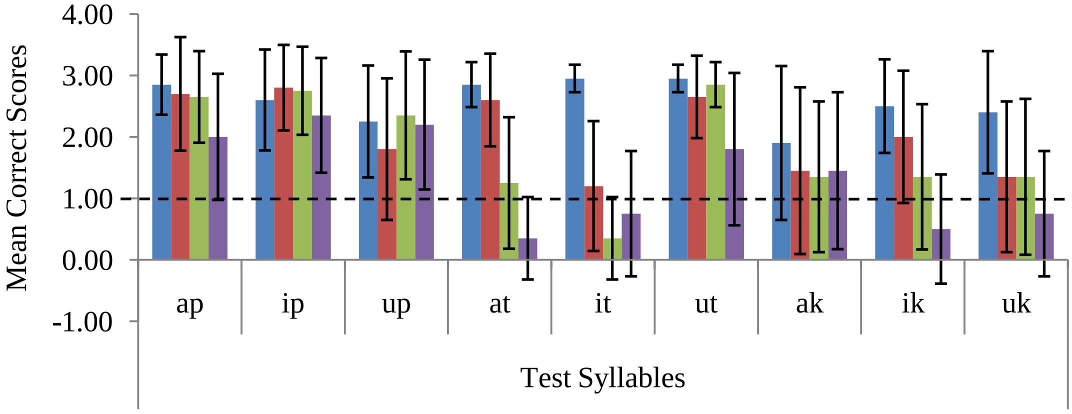


Figure 4.2: Mean and SD of consonant identification scores of normal hearing participants in the four stimulus conditions (VC, VT, VHT, & V), in Quiet condition in VC syllables.
Note: Maximum individual score was 3
.

The horizontal dashed line in the figure demarcates between cues that provide adequate coarticulatory information versus cues that do not. Unlike in CV syllables, in VC syllables most of the cues qualified to be the coarticulatory cues. This was true with most of the test syllables. The exceptions were, V tokens in /at/, /it/, /ik/ and /uk/, and VHT tokens in /it/ alone.

In summary, there were more number of coarticulatory cues identified in VC syllables compared to CV syllables with respect to all the three stop consonants in the three vowel contexts. In all the future data analysis of the present study, only cues identified as providing adequate coarticulatory information in the above two sections have been considered to study the effect of within-group factors and the independent variable. The complete syllable and the stimulus tokens that did not provide adequate coarticulatory information (ones that did not qualify as coarticulatory cues) were left out from the analysis, as the focus was only on the coarticulatory cues.

4.2. Effect of Noise on the Perception of Coarticulatory Cues

To derive the effect of signal to noise ratio on coarticulatory perception, the consonant identification scores obtained in the Quiet condition were compared with that of 10 dB SNR. Identification scores only in those cues which qualified as coarticulatory cues in the section 4.1 were used for the purpose. Only the data from normal group was used for the purpose. The results are reported separately for CV and VC syllables.

4.2.1. Results in CV Syllables

Figure 4.3 shows the mean and SD of consonant identification scores of normal hearing group between two SNRs (Quiet & 10 dB SNR) in the three stimulus tokens that provided adequate coarticulatory information, in CV syllables. The figure shows that the mean identification scores were higher in Quiet condition compared to that in 10 dB SNR in most of the syllables, exception being /pa/ in VT, VHT and V tokens and /ka/ in VT token.

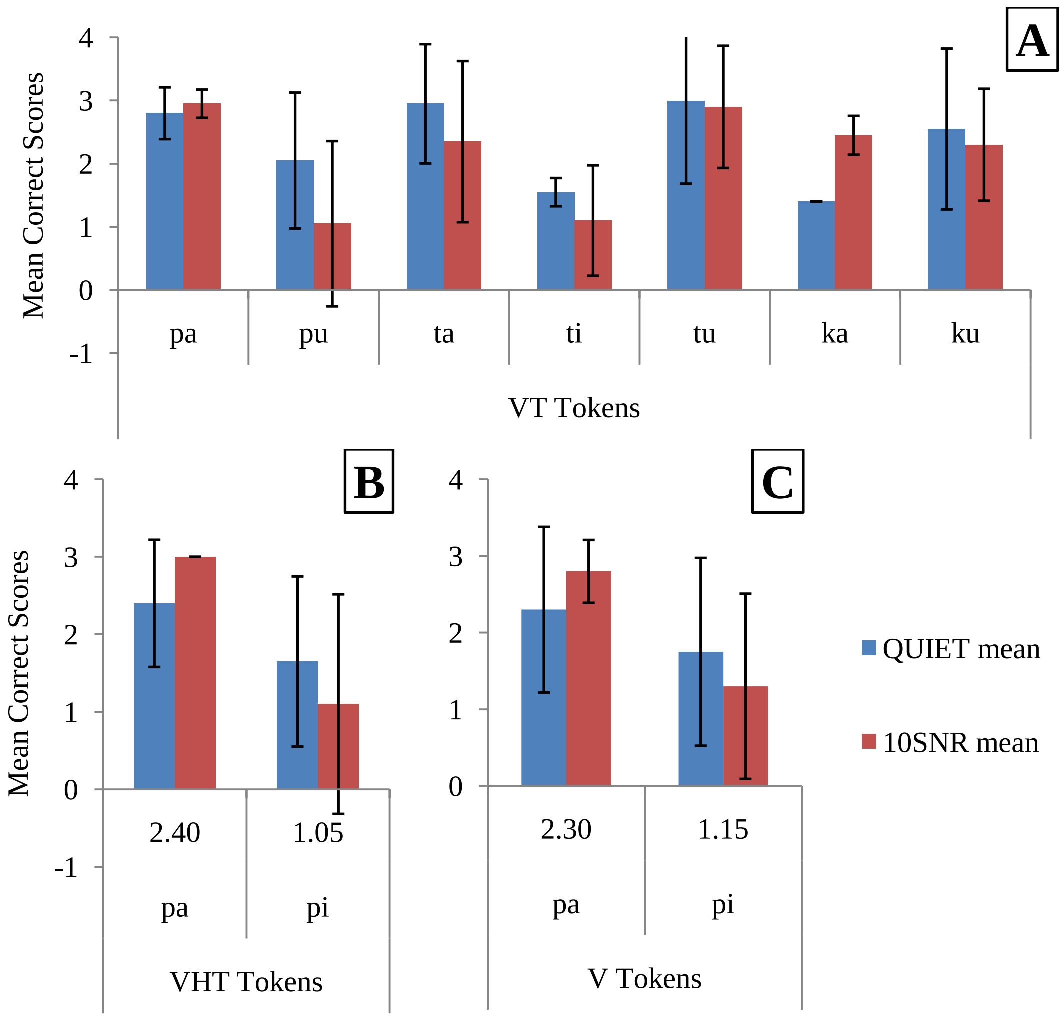


Figure 4.3: Mean and SD of consonant identification scores of normal hearing group between two SNRs (Quiet & 10 dB SNR) in VT (A), VHT (B) and V (C) tokens above chance performance in CV syllables.
Note: Maximum individual score was 3.

To test whether the observed mean differences were significantly different between Quiet and 10 dB SNR conditions, Wilcoxon sign rank test was used. Results of the test are shown in Table 4.1 which shows that there was a significant difference between Quiet and 10 dB SNR in VHT and V tokens of /pa/ and VT tokens of /pu/, /ta/ and /ka/.

Table 4.1: *Results of Wilcoxon test (|z|) in normal hearing group comparing between two SNRs (Quiet & 10 dB SNR) in the three stimulus tokens (VT, VHT, & V) in the CV syllables*

CV	Wilcoxon Test CV QUIET-10 dB SNR		
	VT	**VHT**	**V**
/pa/	1.342	2.585*	1.995*
/pu/	2.187*	1.440	1.246
/ta/	2.489*	-	-
/ti/	1.572	-	-
/tu/	1.414	-	-
/ka/	2.623*	-	-
/ku/	0.779	-	-

Note: * indicates $p < 0.05$

	Quiet > 10 dB SNR		10 dB SNR > Quiet

4.2.2. Results in VC Syllables

Figure 4.4 shows the mean and SD of consonant identification scores of normal hearing group between two SNRs (Quiet & 10 dB SNR) in the three stimulus tokens that provided adequate coarticulatory information, in VC syllables. The figure shows the mean identification scores were higher in Quiet condition compared to that in 10 dB SNR in most of the syllables, exception being /at/, /it/, /ut/ and /ak/ in VT token and /ap/ and /uk/ in V token.

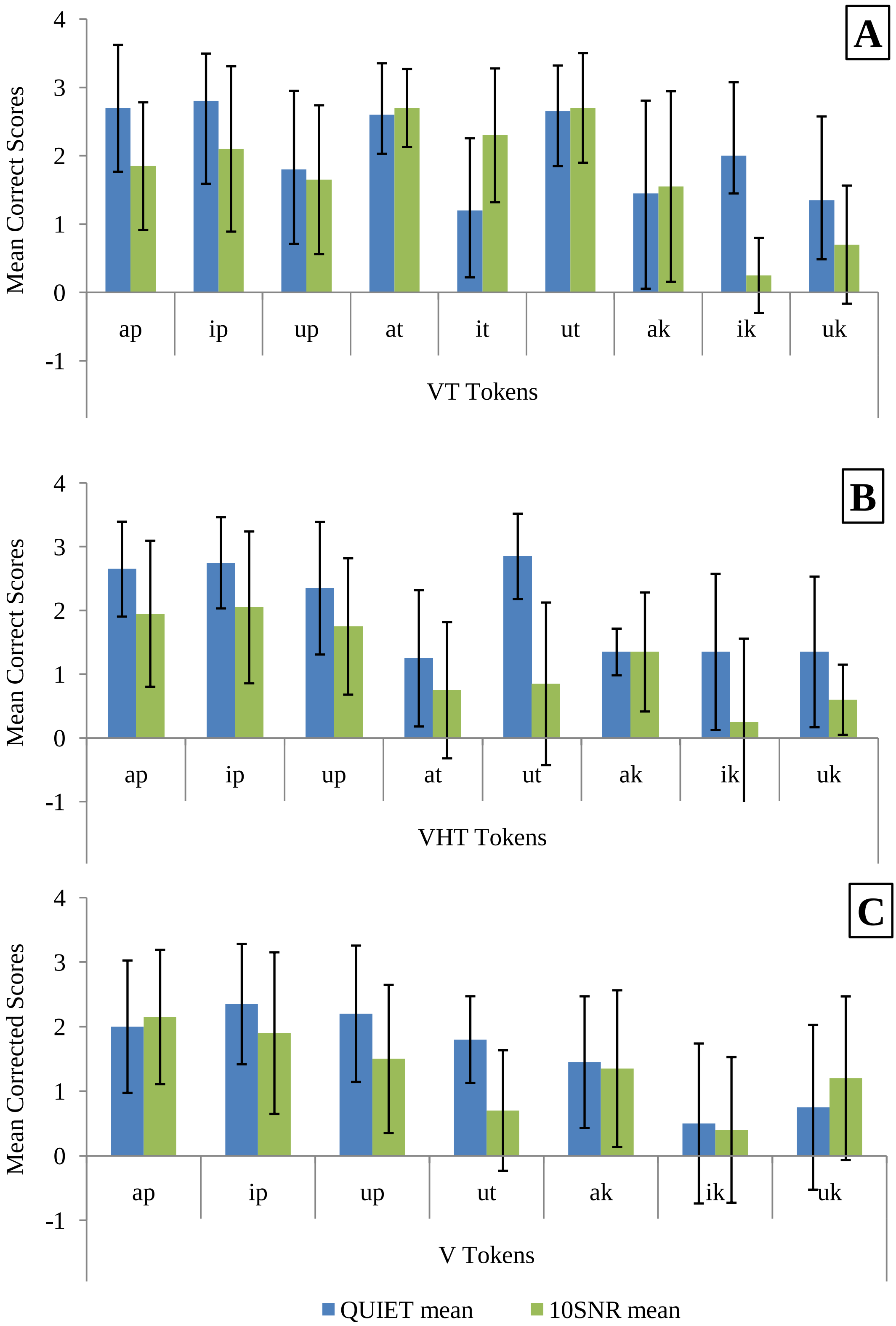


Figure 4.4: Mean and SD of consonant identification scores of normal hearing group between two SNRs (Quiet & 10 dB SNR) in VT (A), VHT (B) and V (C) tokens above chance performance in VC syllables.
Note: Maximum individual score was 3.

To test whether the observed mean differences were significantly different between Quiet and 10 dB SNR conditions, Wilcoxon sign rank test was used. Results of the test are shown in Table 4.2 which shows that there was a significant difference between Quiet and 10 dB SNR in VT token of /ap/, /it/, /ik/, and /uk/, VHT token in /ap/, /ip/, /up/, /ut/, /ik/ and /uk/ and V token in /ut/. In general, the effect of noise seemed to be more in transition based cue than the vowel based cue.

Table 4.2: *Results of Wilcoxon test* **|z|** *in normal hearing group comparing between two SNRs (Quiet & 10 dB SNR) in three stimulus tokens (VT, VHT, & V) in VC syllables*

VC	Wilcoxon Test VC QUIET- 10 dB SNR		
	VT	VHT	V
/ap/	2.984*	2.392*	0.464
/ip/	1.876	2.074*	1.153
/up/	0.431	1.997*	1.775
/at/	0.284	1.523	-
/it/	2.811*	-	-
/ut/	0.276	3.816*	2.716*
/ak/	0.198	0.158	0.215
/ik/	3.674*	2.896*	-
/uk/	2.105*	2.359*	-

Note: * indicates $p < 0.05$

(grey)	Quiet > 10 dB SNR	(yellow)	10 dB SNR > Quiet

4.3. Effect of Carryover versus Anticipatory Coarticulatory Cues

To compare between anticipatory and carryover coarticulatory cues, the identification scores in the CV tokens were compared with that of the corresponding VC tokens. Here again, only the truncated tokens were considered while the complete syllable was ignored as the focus was in coarticulatory cues. The results are discussed separately for Quiet and 10 dB SNR condition.

4.3.1. Results in Quiet Condition

Figure 4.5(A) shows the mean and SD of the consonant identification scores in the three truncated tokens of CV and VC syllables in Quiet condition.

In the Figure 4.5(A), consonants /p/ in vowel context /i/ and /k/ in vowel context /a/ showed higher mean scores for anticipatory condition (VC) in all the three tokens (VT, VHT & V) compared to carryover condition (CV). On the other hand, /t/ in the context of /i/ showed higher mean scores for carry-over condition compared to anticipatory condition. However, no such uniform trend was found in the other syllables.

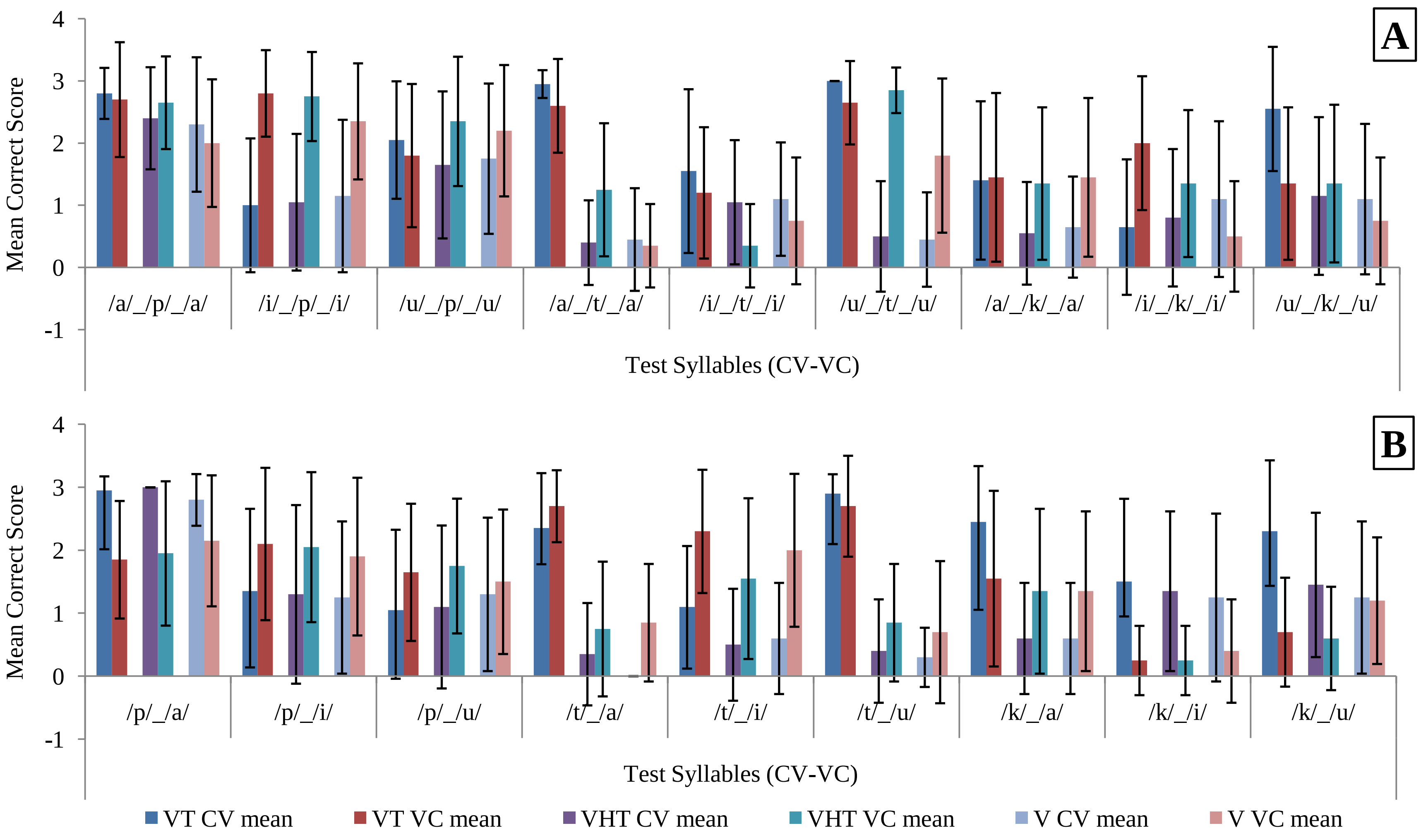


Figure 4.5: Mean and SD of consonant identification scores of normal hearing participants in carryover and anticipatory conditions, in the three stimulus tokens (VT, VHT, & V) in QUIET (A) condition and 10 dB SNR (B) condition.
Note: Maximum individual score was 3
.

Table 4.3: *Results of Wilcoxon test |z| in normal hearing group between VC and CV in three stimulus tokens (VT, VHT, & V) in Quiet condition*

CV-VC	**Wilcoxon Test QUIET Condition**		
	VT	**VHT**	**V**
/ap-pa/	0.108	1.099	0.966
/ip-pi/	3.671**	3.564**	2.656*
/up-pu/	1.097	1.877	0.922
/at-ta/	1.897	2.676*	0.343
/it-ti/	0.918	2.359*	0.901
/ut-tu/	2.070*	3.926*	3.017*
/ak-ka/	0.212	2.071*	2.385*
/ik-ki/	2.913*	2.500*	1.702
/uk-ku/	3.002*	0.477	1.035

Note: * indicates p < 0.05, ** indicates p < 0.01

(grey)	CV > VC	(yellow)	CV < VC

Wilcoxon sign rank test was used to compare the mean identification scores between carryover and anticipatory cues in Quiet condition. This is given in Table 4.3. Results showed that significant difference between CV and VC tokens were seen more in the transition cues (VT and VHT) compared to vowel cue. Within the transition cues, the difference was seen more in VHT than VT tokens. Specifically, there was a significant difference in the identification of /p/ in the context of /i/, /t/ in the context of /u/, in all the three tokens. Similar trend was seen for /k/ in the context of /a/ in VHT and V tokens and in the context of /i/ in VT and VHT tokens.

4.3.2. Results in 10 dB SNR Condition

In the Figure 4.5(B), consonants /p/ in the context of /i/ and /u/, and /t/ in the context of /a/ and /i/ showed higher mean scores in anticipatory condition compared to carry-over condition in all the three tokens (VT, VHT & V). In case of /p/ in the context of /a/ and /k/ in the context of /i/ and /u/ showed better performance in carry-over condition compared to anticipatory condition in all the three tokens (VT, VHT & V). But in case of /t/ in vowel context /u/ and /k/ in vowel context /a/ showed better performance in carry-over condition for VT token and vice versa for VHT and V tokens.

Table 4.4: *Results of Wilcoxon test |z| comparing between VC and CV in normal hearing group, in the three stimulus tokens (VT, VHT, & V) in 10 dB SNR condition*

CV-VC	**Wilcoxon Test 10 dB SNR Condition**		
	VT	**VHT**	**V**
/ap-pa/	3.397**	2.969*	2.228*
/ip-pi/	1.734	1.759	1.462
/up-pu/	1.162	1.554	0.551
/at-ta/	1.461	1.257	3.002*
/it-ti/	2.925*	2.446*	2.639*
/ut-tu/	1.069	1.508	1.192
/ak-ka/	2.250*	1.956	1.957*
/ik-ki/	3.087*	2.889*	2.264*
/uk-ku/	3.386**	2.623*	0.192

Note: * indicates $p < 0.05$, ** indicates $p < 0.01$

(grey)	CV > VC	(yellow)	CV < VC

Wilcoxon sign rank test was used to compare the identification scores between carryover and anticipatory condition in 10 dB SNR condition. This is given in Table 4.4.

Among the tokens that had higher mean scores in anticipatory condition compared to carryover condition, significant difference was present in consonant /t/ in vowel context /i/ in all the three tokens and in consonant /t/ and /k/ in vowel context /a/ in V token. On the other hand, in the tokens that had higher mean scores in carry-over condition compared to anticipatory condition, significant difference was present in consonant /p/ in vowel context /a/ and /k/ in vowel context /i/ in all the three tokens. Also in consonant /k/ in vowel context /a/ in VT token and in vowel context /u/ in VT and VHT tokens.

4.4. Effect of Group on Coarticulatory Perception in Quiet and 10dBSNR

Effect of group was determined by comparing the consonant identification scores between normal and SNHL groups. The consonant identification scores of only those tokens which had adequate coarticulatory cues (as identified in section 4.1) were compared between the two groups. Results are reported separately for CV and VC tokens and separately for quiet and 10dBSNR conditions.

4.4.1. Results of CV Syllables

Figure 4.6 (A & B) shows the mean and SD of identification scores of normal and SNHL group obtained in Quiet (A) and 10dBSNR (B) condition in CV syllables. In most of the tokens of quiet condition, mean scores were higher in normal group compared to SNHL group, except for V token of /pa/ and VT of /ka/. Similarly in most of the cases of 10 dB SNR, mean identification scores were higher in the control group, except for V token of /pa/ and /pu/, and VT token of /tu/. The mean differences were compared for statistical significance using Mann-Whitney U test (Table 4.5) and the results of Quiet condition showed significant difference in consonants /pa/, /pu/ and /ku/ in VT token. On the other hand, in 10 dB SNR, significant differences were found only in VT token of /ta/ and /ti/. In all the five tokens, higher mean value was present in the normal hearing group.

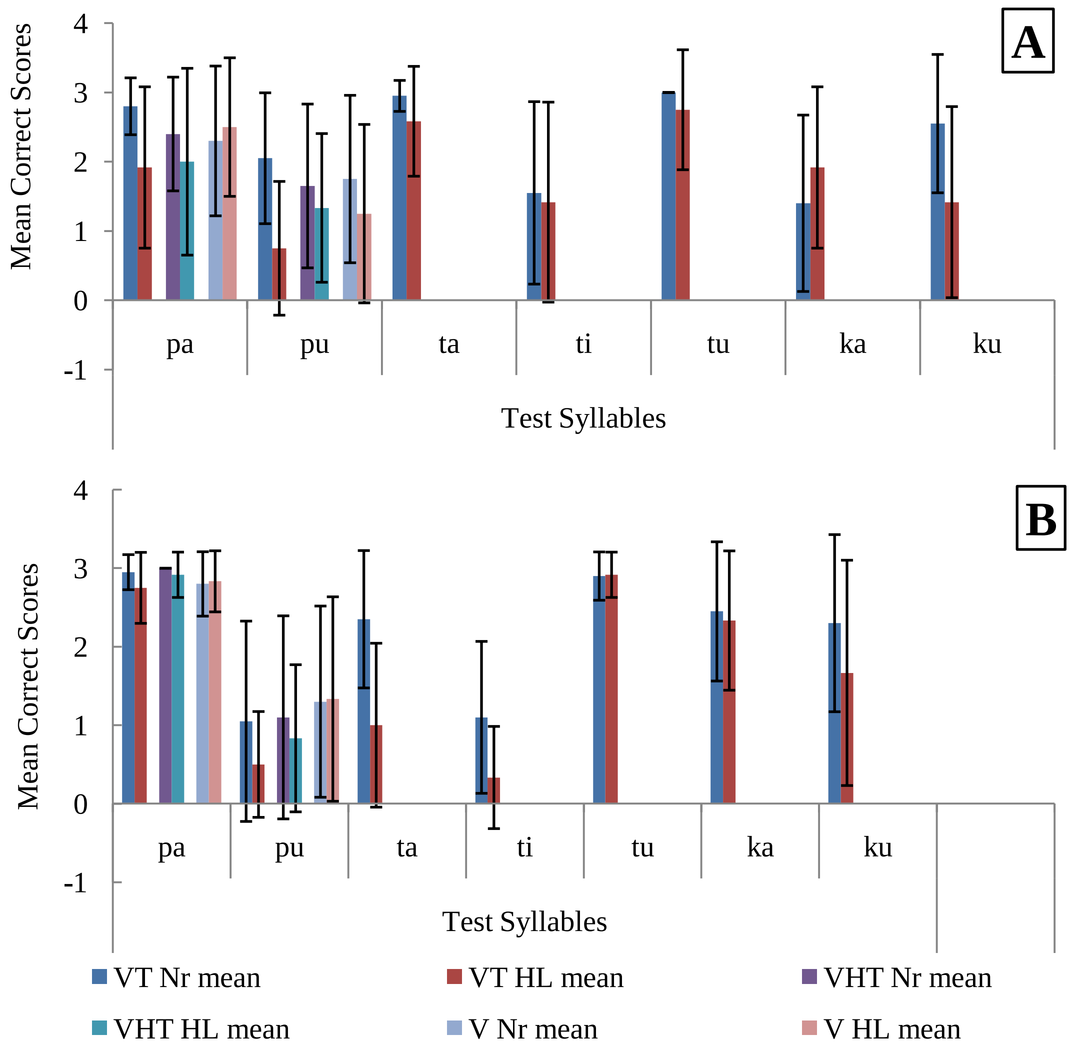


Figure 4.6: Mean and SD of consonant identification scores of normal hearing and SNHL participants in three stimulus tokens above chance level in CV syllables in Quiet condition (A) and 10 dB SNR (B) condition.

Table 4.5: *Results of Mann Whitney U test |z| in Quiet and 10 dB SNR condition comparing normal hearing and SNHL participants in the three stimulus tokens (VT, VHT, & V), in the CV syllables*

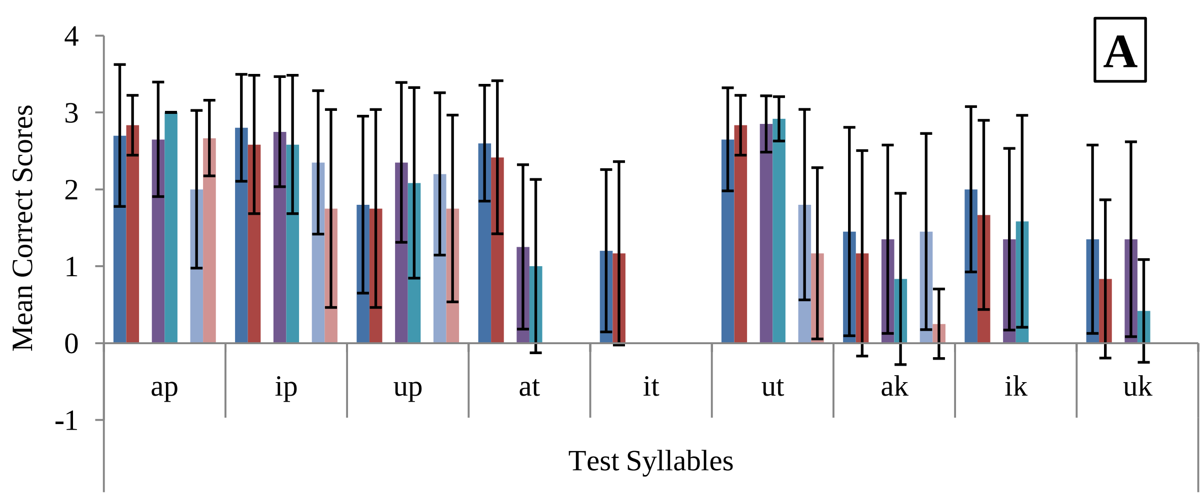
SNR	**CV**	**Mann Whitney U Test QUIET**		
		VT	**VHT**	**V**
QUIET	**/pa/**	2.500*	0.616	0.570
	/pu/	3.139*	0.784	1.134
	/ta/	1.694	-	-
	/ti/	0.331	-	-
	/tu/	1.291	-	-
	/ka/	1.113	-	-
	/ku/	2.593*	-	-
10 dB SNR	**/pa/**	1.630	1.291	0.230
	/pu/	0.928	0.461	0.000
	/ta/	3.214*	-	-
	/ti/	2.388*	-	-
	/tu/	0.154	-	-
	/ka/	0.428	-	-
	/ku/	1.146	-	-

Note: * indicates $p < 0.05$

(grey)	Normal group > SNHL	(yellow)	Normal group < SNHL

4.4.2. Results of VC Syllables

Figure 4.7 (A & B) shows the mean and SD of identification scores of normal and SNHL group obtained in Quiet (A) and 10dBSNR (B) condition in VC syllables. In most of the tokens of quiet condition, mean scores were higher in normal group compared to SNHL group, except for all three tokens of /ap/, VT and VHT tokens of /ut/ and, only VHT token of /ik/. However in most of the cases of 10 dB SNR, mean identification scores were higher in SNHL group compared to normal group, except for all three tokens of /ak/, VT and VHT tokens of /up/ and /uk/, only VT token of /ik/ and only VHT token of /ip/. The mean differences were compared for statistical significance using Mann-Whitney U test (Table 4.6) and the results of Quiet condition showed significant difference in only VHT token of /uk/ and V token of /ak/. In these two tokens, higher mean scores were present in the normal group. On the other hand, in 10 dB SNR, significant differences were found only in VHT token of /ap/, wherein there was higher mean score in the SNHL group.



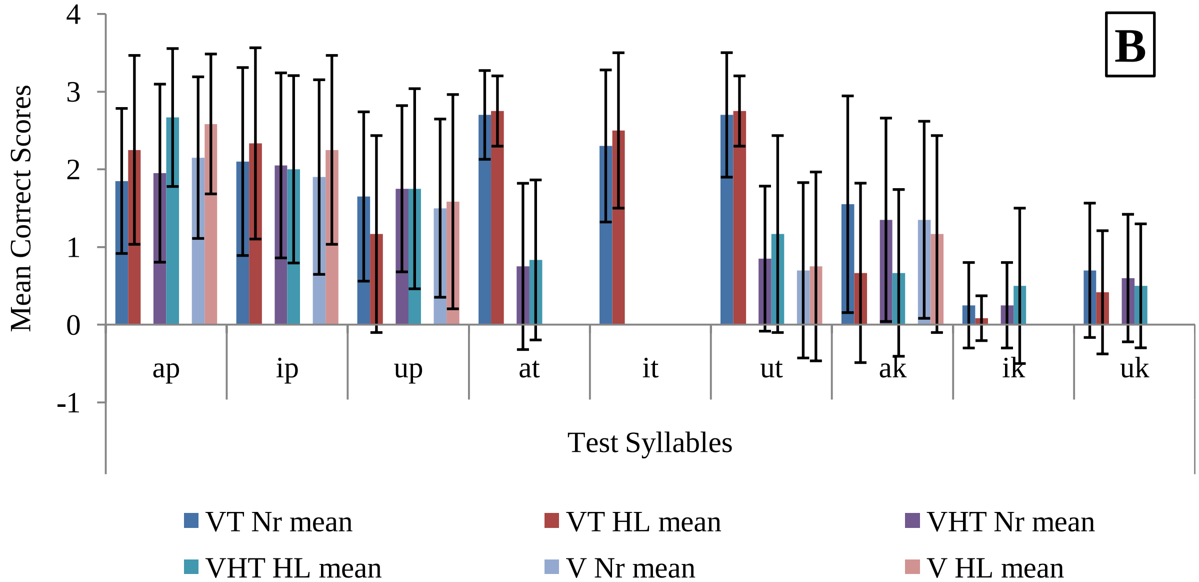


Figure 4.7: Mean and SD of consonant identification scores of normal hearing and SNHL participants in three stimulus tokens above chance level in CV syllables in Quiet condition (A) and 10 dB SNR (B) condition.

Table 4.6: *Results of Mann Whitney U test |z| comparing normal hearing and SNHL participants in the three stimulus conditions (VT, VHT, & V), in Quiet and 10 dB SNR condition in VC syllables*

SNR	**VC**	**QUIET**		
		VT	**VHT**	**V**
QUIET	**/ap/**	0.407	1.852	1.865
	/ip/	1.079	0.687	1.339
	/up/	0.020	0.539	1.143
	/at/	0.358	-	-
	/it/	0.162	-	-
	/ut/	0.648	0.543	1.412
	/ak/	0.601	1.233	2.762*
	/ik/	0.776	0.446	-
	/uk/	1.163	2.034*	-
10 dB SNR	**/ap/**	1.341	2.001*	1.367
	/ip/	0.855	0.084	0.981
	/up/	1.212	0.101	0.202
	/at/	0.077	-	-
	/it/	0.902	-	-
	/ut/	0.515	0.499	0.000
	/ak/	1.653	1.532	0.455
	/ik/	0.895	-	-
	/uk/	1.128	-	-

Note: * indicates $p < 0.05$

(grey)	Normal group > SNHL	(yellow)	Normal group < SNHL

Chapter 5

DISCUSSION

Coarticulation leads to the spread of acoustic characteristics of one phoneme over the other. Perceptually it is known to have a facilitative function, wherein certain phonemes in the context of others are identified better than otherwise. The present study aimed at probing into one such facilitative function of the coarticulation i.e. perceptual utility of coarticulatory cues in the presence of noise. Consonants are weak in their amplitudes but, significantly contribute for intelligibility of speech. Vowels on the other hand do not directly contribute much for the intelligibility of speech. However, in the coarticulatory context with consonants they provide boost to the consonantal cues leading to better perception of consonants. In the presence of noise, consonantal cues tend to get masked and may not be perceived if in isolation. The boost given by the vowels shall be helpful in the perception of the consonantal cues, and thereby facilitate speech perception in noise. With this theme the present study was taken up, and the findings of the present study are discussed under the following headings:

1) Coarticulatory cues for the perception of unvoiced stop consonants in Kannada
2) Relative importance of carryover and anticipatory coarticulatory cues in the perception of stop consonants
3) Perceptual utility of coarticulatory cues in noisy environment

4) Influence of cochlear HL on the perception of stop consonants based on coarticulatory cues

5.1. Coarticulatory Cues for the Perception of Unvoiced Stop Consonants in Kannada

In the present study three unvoiced stops /p, t, k/ of Kannada in the context of /a, i, u/ were taken up to investigate the perceptual cues of coarticulation. The target feature studied was the place of articulation of stop consonants and this was studied in anticipatory as well as carryover coarticulatory contexts.

In the present study coarticulatory cue was defined as a cue that does not occur when the consonant is produced in isolation, but occurs in coarticulatory context. Therefore, with reference to stop consonants, burst would be considered as a primary consonantal cue while transition and adjacent vowel would be considered as coarticulatory cues. As a preliminary step in the present study, attempt was made to identify perceptual utility of coarticulatory cues (full transition with vowel, half transition with vowel & only vowel) for the identification of consonants. For example, if a consonant could be identified just with the adjacent coarticulated vowel, in spite of removing burst and transitions, the vowel was designated as perceptually useful as a coarticulation cue.

5.1.1. Role of Carryover Coarticulatory Cues

The findings of the present study showed that, in carryover coarticulation (derived from CV syllables) primarily full transition with vowel served as useful coarticulatory cues. The exception to this was /pa/ and /pu/, wherein consonant could be identified even from half transition with vowel and only vowel. This means that there was a greatest spread of carryover effect of /p/ in the context of vowels /a/ and /u/.

The better perception of /p/ in the context of /a/ and /u/ can be attributed to feature spreading. That is, /a/ and /u/ being mid and low vowels enhance the perception of bilabial /p/ which has lower burst frequency (Schatz, 1954; Halle, Hughes & Radley, 1957). The finding can also be attributed to the vowel duration in the syllables /pa/ and /pu/. Revoile, Pickett, and Holden (1982) reported as the vowel duration of /a/ increases, listener tend to identify as labials. Vowel /a/ and /u/ being mid and low vowels have got longer duration compared to /i/ (Lindblom, 1963). Therefore, /p/ was probably identified better in the context of longer duration vowels.

Further, compared to /a/ and /u/, in the context of /i/, all the three stop consonants were perceived poorer based on the coarticulatory cues. This is in agreement with the earlier literature (Singh & Black, 1965; Wang & Bilger, 1973; Dubno, Dirks, & Langhofer, 1982). This can be attributed to the similarity in the direction of F2 transition across the three consonants. In the context of /i/, all the three consonants had rising F2 and therefore, leading to confusion while identifying the place of articulation. On the other hand direction of coarticulation was different across the consonants in the context of /a/ and /u/.

Of the three consonants, /t/ was perceived better in the context of /i/ based on the full transition with vowel, which can be again attributed to feature spreading. Poor performance in the context of /i/ implies coarticulatory cues do not facilitate perception of place of articulation of stop consonant in the context of /i/. In other words, in the context of /i/ one needs to depend on the primary consonant cue, such as burst or acoustic locus.

5.1.2. Role of Anticipatory Coarticulatory Cues

Unlike in carryover condition, in anticipatory condition, majority of instances, full transition with vowel, half transition with vowel and only vowel helped in the identification of place of articulation of stop consonants. This suggests, that the coarticulatory cues are spread upto vowels in case of VC syllables. Compared to only vowel, transition with vowel cued better for the place of articulation, which suggests the duplex cues are better.

However, the spread of coarticulation was relatively poor in /at/, /it/, /ik/ and /uk/. Of these, coarticulatory spread was present upto the end of transition in /at/, /ik/ and /uk/, while it was present only upto half transition in /it/. This could be because, in /it/ the duration of transition is less compared to all the other syllables. Studies in the literature have also shown that the anticipatory coarticulatory cues provide useful information for the identification of stop consonants (Brady, House & Stevens, 1961; Sharf & Hemeyer, 1972; Massaro, 1973; Ohde & Sharf, 1977; Pols & Schouten, 1978; Sharf & Beiter, 1974).

On comparing across the three consonants, /p/ showed greater anticipatory effect followed by /k/ and /t/ in all the three truncated tokens across the three vowels /a, i, u/. Similar findings was reported by earlier studies (Wang & Bilger, 1973; Modarresi, Sussman, Lindblom, & Burlingame, 2004) for the stop consonants. Bilabials and velars show more of anticipatory effect, whereas alveolars show more of carryover effect. However, in /t/ anticipatory effect was seen across all the truncated tokens, which was highest in the context of vowel /u/ followed by /a/ and /i/. The prominent acoustical difference between these tokens was in terms of greater contrast effect in terms of F2 in /ut/ compared to /at/ and /it/. Additionally there was also decrease in transition duration in the same.

5.2. Relative Importance of Carryover and Anticipatory Coarticulatory Cues in the Perception of Stop Consonants

On comparing anticipatory and carryover coarticulation, more number of coarticulatory cues were perceptually useful than that in carryover coarticulation. This suggests that, anticipatory coarticulation are stronger and are perceptually more important than the carryover coarticulation cues. This relative advantage of anticipatory cues is true for all three places of articulation and in all three vowel contexts except for /p/ in the context of /a/.

Statistical comparison of anticipatory and carryover cues revealed that, majority of difference between the two types of cues was in the token that had half transition with vowel. This means that the spread of coarticulation was beyond transition in anticipatory, but only upto first half of the transition in carryover coarticulation. Further, the presence of significance of difference itself means that, anticipatory cues were more robust compared to carryover cues. Exception to this general trend was /t/ and /k/ in the /u/ context in VT, and /t/ in the /i/ context in VHT, wherein carryover coarticulation was better than anticipatory coarticulation. The findings of the present study are in agreement with the studies in the literature (Brady, House & Stevens, 1961; Sharf & Hemeyer, 1972; Massaro, 1973; Ohde & Sharf, 1977; Pols & Schouten, 1978; Sharf & Beiter, 1974).

Ohde and Sharf, (1977) recorded CV and VC syllables using the stops /p, t, k, b, d, g/ and vowels /a, i, u/ to study the order effect of acoustical segments in identifying the stops and vowels. The obtained CV and VC syllables were separated into aperiodic, aperiodic+ vocalic transition, vocalic transition and vocalic transition+ vowel segments through tape deletion method. Results revealed that consonant identification scores were greater for VC syllables compared to CV syllables in vocalic transition+ vowel segments, but no significant difference were seen in other three tape deletions. The present study used slightly different method compared to Ohde and Sharf (1977). Method in the present study helped characterizing the spread of coarticulation more precisely within transition (through comparing between full transition and half transition conditions). compared to the study of Odhe and Sharf.

Pols and Schouten (1978) investigated the identification of the deleted initial and final consonant with and without 300 ms noise burst stimuli preceding (CV) or following (VC) the stimuli. The consonants used in the study were /p, t, k, b, d/+ /i, u, a/+ /t/ in the initial position and /t/+ /i, u, a/+ /p, t, k/ in the final position. Results revealed identification of consonant with VC transition is better than CV transition. Because, when a burst is truncated from the CV or VC stimuli it results in click like sensation. When the click is in the beginning of the transition, due to its abrupt beginning the scores of the CV transition reduces compared to VC transition as the click occurs at the end results in burst like perception. When noise burst were introduced near the click in case of VC transition, identification of consonants did not very much with or without noise burst. But in case of CV transition, identification of consonants improved with noise. This may be due to masking effect of the noise burst over the click like sensation. Also they report response bias towards the identification of /p/ for the deleted consonants in both CV and VC conditions even in the presence of noise. This may be due to perception of the click, like burst of consonant /p/. This suggests that in anticipatory coarticulatory condition, the perception is less dependent on the consonantal cues.

Sharf and Hemeyer (1971) listed three models for the identification of consonants better in VC formant transition than CV: transition direction, transition dependence, and transition sufficiency model. In the *transition direction model*, it depends on the direction of change in frequency and rate of change in frequency. The direction of change in frequency in the CV transition reaches or leans upon the vowel steady state which is the terminal frequency of the transition, whereas in VC transition it reaches the consonant.

This terminal frequency provides more information on the vowel in CV and final consonant in VC (Brady, House & Stevens, 1961). In the *transition dependency model*, the information provided by the vowel transition for consonant identification is based on the information provided by the burst frequency of the stops. The burst frequencies of the initial stops are more important cue for perception than the burst frequency of the final stops. House et al. (1965) support this model by reporting; at various signal-to-noise ratios initial stops are better perceived than final stops. In American English pronunciation the final stops bursts are unreleased, as there is insufficient pronunciation of the final consonant, the listener depend more on VC transition than CV transition. In the *transition sufficiency model,* VC formant transition provides better information on place of articulation than CV formant transition. In CV transition, part of the transition is represented in the burst duration of the consonant which leads to difficulty of stops perception when the burst frequency is truncated in the CV than VC. This leads to difficulty in making decision for CV transition than VC transition. This may be due to phonetic assimilation where the effect of forward coarticulation is greater than backward coarticulation.

5.3. Perceptual utility of Coarticulatory Cues in Noisy Environment

Comparisons of perceptual utility of coarticulation cues between Quiet and 10 dB SNR showed that the effect of noise was bidirectional. In most instances, perception was poorer in 10 dB SNR, compared to Quiet condition. This is because consonantal cues are low in their amplitude and therefore get easily masked in the presence of noise. In such conditions perception is based only on the secondary coarticulatory cues and therefore, decrease in scores is not a surprising finding. The decrease in the perception in the presence of noise was primarily seen in full transition with vowel and half transition with vowel condition. This suggests that the addition of the noise was masking some amount of the transition. Therefore, one will have to depend more on the second half of the transition and the vowel for the coarticulatory information. Considering that the spread of coarticulation is lesser in carryover compared to anticipatory, one can expect that the carryover coarticulation will not be useful in the presence of noise. Anticipatory coarticulation, because is spread upto vowel shall be useful in the presence of noise.

However, there were also instances were perception based on coarticulatory cues was better in the presence of noise compared to Quiet. Although the exact reason is not known, it is speculated that it is because of perceptual restoration. This is supported by the findings of earlier studies (Clarke, Gaudrain, Chatterjee, & Başkent, 2014; Saija, Akyürek, Andringa, & Başkent, 2014). Pols and Schouten (1978) investigated the identification of the deleted initial and final consonant with and without 300 ms noise burst stimuli preceding (CV) or following (VC) the stimuli. Results showed, increase in the identification of consonants with CV transitions. Probably the addition of noise in the deleted regions of syllables gives a burst like quality and therefore, is leading to better perception.

Furthermore, comparison of anticipatory and carryover coarticulation in the presence of noise showed, that the anticipatory coarticulation is better than carryover coarticulation except for /p/ in /a/ context and, /k/ in the /i/ and /u/ context. This supported or earlier presumption based on the spread of coarticulation that, anticipatory coarticulation would be better than carryover coarticulation in the presence of noise.

Future research may consider additional signal-to-noise ratios (SNRs) and compare the perception of coarticulatory cues across SNRs as the information in the present study is restricted to one SNR. Gordon-Salant & Wightman (1983) reported that as the SNR improves, the identification scores of the consonants improve reaching to the level of Quiet condition above 10 dB SNR (Gordon-Salant & Wightman, 1983). Therefore SNR below 10 dB might be a more realistic picture of the effect of noise on the perception based on coarticulatory cues.

In several tokens there was no significant difference between Quiet and 10 dB SNR conditions. This suggests that coarticulatory cues are helpful for the speech perception of noise.

5.4. Influence of Cochlear Hearing Loss on the Perception of Stop consonants based on Coarticulatory Cues

Coarticulatory perception was expected to be reduced in individuals with cochlear hearing loss compared to normal hearing individuals, due to the compromised spectral and temporal resolution characteristics of cochlear pathology. Results showed significantly poorer perception of place of articulation of stop consonants in individuals with SNHL compared to normals. In the carryover coarticulation, this was seen primarily in the speech token having voice transition with vowel. The reduced perception can be attributed to reduce frequency selectivity (Moore, & Glasberg, 1986; Moore, 2001; Kluk & Moore, 2005) compressive nonlinearity (Moore, 1998), reduced temporal resolution (Viemeister, 1979; Hopkins & Moore, 2009). The finding that, the group difference was evident only in VT can be attributed to the spread of coarticulation. Because in carryover coarticulation, the spread of coarticulation was restricted to the first half of the transition, the group difference was seen only in the VT token. The absence of group difference in the VHT and V tokens could be because of floor effect. However, in VC tokens the group differences are seen in VHT and V tokens also.

In the 10 dB SNR of /ap/ syllable, better performance was observed in individuals with cochlear hearing loss. Although the exact reason is not known for this finding, it can be speculated that, it is because of lesser dependence of individuals with cochlear hearing loss on spectral cues. The addition of noise that reduces the spectral contrasts did not affect their perception while it reduced perception in normal hearing individuals who depend on both spectral and temporal cues.

Helfer and Huntley (1991) compared the consonant recognition in older adults with minimal HL to moderate high frequency HL and young adults with normal hearing in the presence of cafeteria noise using nonsense syllables. Across both the groups in place of articulation, bilabials and dentals were most affected in both the groups with better scores in young adults.

To summarize, the perceptual weightage of coarticulatory cues is primarily dependent on whether it is anticipatory or carryover coarticulation. Signal-to-noise ratio and hearing loss, although effects the perception of coarticulatory cues, the effect is again dependent on the direction of coarticulation and consonant vowel combination. Furthermore, coarticulation cues identified in the present study and the influences drawn in terms of their perceptual weightage, cannot be generalized to the languages other than Kannada, unless experimentally tested. This is because, there could be cross linguistic differences in the coarticulatory cues and their perception.

Chapter 6

SUMMERY AND CONCLUSION

The present study aimed at documenting the effect of following three attributes on the perception of stop consonants based on coarticulatory cues;

- Presence of background noise
- Direction of coarticulatory effect (anticipatory & carryover condition)
- Cochlear hearing loss

Three Kannada unvoiced stop consonants /p, t, k/ were combined with three vowels /a, i, u/ to form 9 CV and 9 VC syllables. All the CV and VC syllables were truncated into 3 tokens each i.e., syllables with full transition + vowel (VT), half transition + vowel (VHT) and only vowel (V). These tokens were also mixed with speech shaped noise to obtain a condition with 10 dB SNR along with Quiet condition. Thus, there were 72 tokens each in Quiet and 10 dB SNR. Closed set identification of these syllables was carried out at 40 dB SL in 20 normal hearing and 12 SNHL participants. The stimuli across all the conditions were randomized and presented only to the right ear of all the participants.

The identification scores of all the speech tokens in CV and VC conditions in two SNRs (Quiet & 10 dB SNR) between the two groups were subjected to statistical analysis. It was found that, the raw data did not show normality and also lacks sphericity. Thus, non-parametric tests were used for all the comparisons. Wilcoxon sign rank test was used to compare between quiet versus 10 dB SNR and, anticipatory versus carryover conditions within group. Mann Whitney U test was used for comparison between groups.

The salient results obtained in the results are as follows:

- Full transition with vowels served as the most useful carryover coarticulatory cue. Consonants are best perceived in the context of /a/ and /u/ than in the context of /i/. This could be due to similarities and differences respectively in the direction of F2 transitions across the three places of articulation.
- The spread of anticipatory coarticulatory cues is extended till the sustained portion of the preceding vowel. Bilabials and velars show more anticipatory coarticulation than alveolars.
- Spread of anticipatory coarticulation cues was greater than carryover cues and was significant for consonant perception.
- Addition of noise reduced anticipatory coarticulatory perception more than carryover coarticulatory perception. This effect was bidirectional i.e., in CV syllables it led to both increase and decrease in consonant perception. The reduction in the salient temporal and spectral cues led to decrease in perception, while burst like perception due to the addition of noise led to better perception in a few tokens.
- In SNHL group, perception based on coarticulatory cue was poorer than normal hearing group. This is attributed to reduced frequency selectivity, compressive nonlinearity and reduced temporal resolution in cochlear hearing loss.

Implications of the Study

- The findings of the study help us in better understanding of contribution of coarticulatory cues in adverse listening conditions.
- The findings aid in understanding of the relative importance of carryover cues over anticipatory cues in adverse listening conditions.
- The findings of the study aid in designing better signal processing strategies to improve coarticulatory perception

Future Directions

- The coarticulatory cues perception was studied in 10 dB SNR, so this effect should be studied in lower SNRs to evaluate the effect of higher noise levels on coarticulatory perception which are more representative of natural listening situations.
- The coarticulatory cues in the perception of noise can be studied for different classes of speech sounds like nasals, fricatives, affricates and also in voicing contrasts.
- This can be extended to coarticulatory perception in hearing aid and cochlear implant users.

REFERENCES

Beattie, R. C., Barr, T., &Roup, C. (1997). Normal and hearing-impaired word recognition scores for monosyllabic words in quiet and noise. *British Journal of Audiology*, *31*(3), 153–164.

Bhuvaneshwari, C. S. &Savithri, S. R. (1993). Development of perception of co-articulation.*Unpublished dissertation.*University of Mysore.

Boothroyd, A. (1984). Auditory perception of speech contrasts by subjects with sensorineural hearing loss. *Journal of Speech, Language, and Hearing Research*, *27*(1), 134-144.

Borden, G. J., & Harris, K. S. (1980).*Speech science primer: Physiology, acoustics, and perception of speech*.Baltimore: Williams & Wilkins.

Brady, P. T., House, A. S., & Stevens, K. N. (1961).Perception of sounds characterized by a rapidly changing resonant frequency.*The journal of the acoustical society of America, 33 (10),* 1357-1362.

Clarke, J., Gaudrain, E., Chatterjee, M., &Başkent, D. (2014). T'ain't the way you say it, it's what you say – Perceptual continuity of voice and top–down restoration of speech. *Hearing Research*, *315*, 80–87.http://doi.org/10.1016/j.heares.2014.07.002

Cooper, F. S. (1953). Erratum: Some Experiments on the Perception of Synthetic Speech Sounds. *Journal of The Acoustical Society of America - J ACOUST SOC AMER*, *25*(3). http://doi.org/10.1121/1.1906940

Cooper, F. S., Delattre, P. C., Liberman, A. M., Borst, J. M., and Gerstman, L. J. (1952). Some experiments on the perception of synthetic speech sounds. *The journal of the acoustical society of America, 24 (6),* 597-606.

Cooper, F. S. (1952). Some Experiments on the Perception of Synthetic Speech Sounds.*Journal of the Acoustical Society of America*, *24*(6), 597-606.

Daniloff, R. G., & Moll, K. L. (1968).Coarticulation of lip rounding.*Journal of Speech and Hearing Research*, *11*(4), 701-721.

Dubno, J. R., Dirks, D. D., &Langhofer, L. R. (1982).Evaluation of hearing-impaired listeners using a nonsense-syllable test II.Syllable recognition and consonant confusion patterns.*Journal of Speech, Language, and Hearing Research*, *25*(1), 141–148.

Dubno, J. R., Dirks, D. D., & Schaefer, A. B. (1987). Effects of hearing loss on utilization of short-duration spectral cues in stop consonant recognition. *The Journal of the Acoustical Society of America, 81*(6), 1940–1947.

Dubno, J. R., Dirks, D. D., & Schaefer, A. B. (1987). Effects of hearing loss on utilization of short-duration spectral cues in stop consonant recognition. *Journal of the Acoustical Society of America*, *81*(6), 1940-1947.

Fernandes, T., Ventura, P., & Kolinsky, R. (2007).Statistical information and coarticulation as cues to word boundaries: A matter of signal quality.*Attention Perception & Psychophysics*, *69*(6), 856-864.

Halle, M., Hughes, G. W., &Radley, J. P. (1957).Acoustic properties of stop consonants. *Journal of the Acoustical Society of America, 29*(1), 107-116.

Hedrick, M. S., &Jesteadt, W. (1996).Effect of relative amplitude, presentation level, and vowel duration on perception of voiceless stop consonants by normal and hearing-impaired listeners.*The Journal of the Acoustical Society of America*, *100*(5), 3398–3407.

Hedrick, M. S., Schulte, L., &Jesteadt, W. (1995).Effect of relative amplitude on perception of voiceless stop consonants by listeners with normal and impaired hearing.*The Journal of the Acoustical Society of America, 98(3),* 1292-1303.

Hopkins, K. & Moore, B. C. J. (2009).The contribution of temporal fine structure to the intelligibility of speech in steady and modulated noise.*Journal of the Acoustical Society of America, 125*(1), 442-446.

House, A. S., Williams, C. E., Hecker, H. L. &Kryter, K. D. (1965). Articulation testing methods: Consonantal differentiation with a closed-response set. *The journal of the acoustical society of America, 37,* 158-166.

Keni, R. D., Wiley, T. L., &Strennen, M. L. (1979).Consonants discrimination as function of presentation level.*Audiology*, *18*, 212-224.

Kluk, K. & Moore, B. C. G. (2005). Factors affecting psychophysical tuning curve for hearing impaired subjects with high frequency dead regions. *Hearing Research, 200(1-2),* 115-131.

Kuhnert, B., & Nolan F. (1999).The origin of coarticulation. In W. J. Hardcastle, & N. Hewlett (Eds.), *Coarticulation - Theory, Data and Techniques* (p. 7). New York, United States of America: Cambridge University Press.

LaRiviere, C., Winitz, H., &Herriman, E. (1975).Vocalic transition in the perception of voiceless initial stops.*The journal of the acoustical society of America, 57,* 470-475.

Lehiste, I., &Shockey, L. (1972).Onthe perception of coarticulation effects in English VCV syllables. *Journal of speech andhearing research*, *15*, 500-506.

Liberman, A. M., Delattre, P. C., & Cooper, F. S. (1958). Some cues for the distinction between voiced and voiceless stops in initial position. *Language and speech*, *1*(3), 153-167.

Lindblom, B. (1963). Spectrographic study of vowel reduction.*The Journal of the Acoustical Society of America*, *35*(11): 1773–1781.

Massaro, D. W. (1973). A comparison of forward versus backward recognition masking.*Journal of experimental psychology, 100,* 434-436.

Modarresi, G., Sussman, H., Lindblom, B., & Burlingame, E. (2004).An acoustic analysis of the bidirectionality of coarticulation in VCV utterances.*Journal of Phonetics*, *32*(3), 291–312. http://doi.org/10.1016/j.wocn.2003.11.002

Moore, B. C. J. (1998). *Cochlear hearing loss.* Whurr publishers: London.

Moore, B. C. J. (2001). Dead regions in the cochlea: diagnosis, perceptual consequences, and implications for the fitting of hearing aid. *Trends in amplification, 5,* 1-34.

Moore, B. C. J. & Glasberg, B. R. (1986). The relationship between frequency selectivity and frequency discrimination for subjects with unilateral and bilateral cochlear impairments, in *Auditory Frequency Selectivity*, B.C.J. Moore and R.D. Patterson, ed., Plenum, New York.

Ohde, R. N., &Sharf, D. J. (1977).Order effect of acoustic segments of VC and CV syllables on stop and vowel identification.*Journal of speech and hearing research, 20*(3), 543-554.

Sharf, D. J. and Odhe, R. N. (1981). Physiological, acoustic and perceptual aspects of coarticulation: implications for the remediation of articulatory disorders. Edited by Lass, N. J. (1981) in *Speech and language: advances in basic research and practice.* Academic press, inc.: New York. Volume 5. Page no. 153-239.

Ohde, R. N., & Stevens, K. N. (1983).Effect of burst amplitude on the perception of stop consonant place of articulation.*Journal of the Acoustical Society of America*, *74(3)*, 706-714.

Ohman, S. E. (1966). Coarticulation in VCV Utterances: Spectrographic Measurements. *Journal of the Acoustical Society of America*, *39*, 151-168.

Pickett, J. M. (1999). *The acoustics of speech communication: Fundamentals, speech perception theory, and technology*.Boston, United States of America: Allyn and Bacon.

Pols, L. C. W., & Schouten, M. E. H. (1978).Identification of deleted consonants.*The Journal of the Acoustical Society of America, 64(5)*, 1333-1337.

Revoile, S., Pickett, J. M., & Holden, L. D. (1982).Acosutic cues to final stop voicing for impaired- and normal-hearing listeners. *The Journal of the Acoustical Society of America*, *72*(4), 1145–1154.

Revoile, Pickett,Holden-Pitt,Talkin,& Brandt (1987). Burst and transition cues to voicing perception for spoken initial stops by impaired and normal hearing listeners. *Journal of Speech and Hearing Research,30* , 3-12.

Revoile,S.G.,Pickett,J.M.& Kozma-Spytek,L. (1991). Spectral cues to perception of/d, n, 1/by normal and impaired hearing listeners. *Journal of Acoustical Society of America, 90*, 787-798.

Saija, J. D., Akyürek, E. G., Andringa, T. C., &Başkent, D. (2014). Perceptual Restoration of Degraded Speech Is Preserved with Advancing Age. *Journal of the Association for Research in Otolaryngology*, *15*(1), 139–148. http://doi.org/10.1007/s10162-013-0422-z

Savithri, Pushpavathi& Rajeev (1994). Transition duration of plosives: developmental data from speech production in children. *Journal of Acoustical Society of India, 23*, 19-27.

Schatz, C. D. (1954). The role of context in the perception of stops.*Language, 30(1)*, 47-56.

Sharf, D. J., &Beiter, R. C. (1974). Identification of Consonants from Formant Transitions Presented Forward and Backward. *Language and Speech*, *17*(2), 110–118. http://doi.org/10.1177/002383097401700202

Sharf, D. J. &Hemeyer, T. (1972).Identification of place of consonant articulation from vowel formant transition.*The Journal of the Acoustical Society of America, 51(2)*, 652- 658.

Sharf, D. J. &Ostreicher, H. J. (1973). Effect of forward and backward coarticulation on the identification of speech sounds.*Language Speech, 16*, 196-206.

Sharf, D.J., &Ohde, R.N. (1981).Physiologic, acoustic and perceptual aspects of coarticulation.In Lass, N J (Ed.), *Speech and Language* (pp. 154-247). New York: Academic press.

Turner, C. W., & Robb, M. P. (1987).Audibility and recognition of stop consonants in normal and hearing-impaired subjects. *The Journal of the Acoustical Society of America*, *81*(5), 1566-1573.

Van Tasell, D. J., Hagen, L. T., Koblas, L. L., &Penner, S. G. (1982). Perception of short-term spectral cues for the stop consonants place by normal and hearing-impaired subjects.*The Journal of the Acoustical Society of America*, *72(6*), 1771-1780.

Viemeister, N. F. (1979). Temporal modulation transfer function based upon modulation thresholds. *The Journal of the Acoustical Society of America*,*66*(5),1364-1380.

Wang, M. D. & Bilger, R. C. (1973). Consonant confusion in noise: a study of perceptual features. *The journal of the acoustical society of America, 54(5),* 1248-1266.

Wang, William, S-Y. (1959). Transition and release as perceptual cues for final plosives. *Journal of speech and hearing research*, *2(1)*, 66-73.

Appendix 1

Table 1: Demographic and Audiometric details of participants in group I (normal hearing participants)

Sl. No	Age (yrs)	Gender	PTA (dB HL)	Tympanogram Type	Acoustic reflex (present at n Hz)	DPOAE (present/ absent)
1	42	M	8.75	A	500, 1k, 2k, 4k	present
2	46	M	15.00	A	500, 1k, 2k	present
3	46	M	15.00	A	500, 1k, 2k	present
4	42	M	13.75	A	500, 1k, 2k	present
5	39	M	13.75	A	500, 1k, 2k	present
6	36	M	10.00	A	500, 1k, 2k	present
7	36	M	5.00	A	500, 1k, 2k	present
8	40	M	13.75	A	500, 1k, 2k	present
9	32	M	12.50	A	500, 1k, 2k	present
10	33	M	11.25	A	500, 1k, 2k	present
11	29	M	8.75	A	500, 1k, 2k	Present
12	29	M	12.50	A	500, 1k, 2k	Present
13	29	M	11.25	A	500, 1k, 2k	Present
14	26	M	12.50	A	500, 1k, 2k, 4k	Present
15	25	M	5.00	A	500, 1k, 2k	Present
16	25	M	2.50	A	500, 1k, 2k, 4k	Present
17	24	F	3.75	A	500, 1k, 2k, 4k	Present
18	24	F	3.75	A	500, 1k, 2k, 4k	Present
19	37	F	5.00	A	500, 1k, 2k, 4k	Present
20	32	F	15.00	A	500, 1k, 2k	Present

Table 2: Demographic and Audiometric details of participants in group II (SNHL participants)

Sl. No	Age (in years)	Gen-der	PTA (dB HL)	Tympanogram Type	Acoustic reflex (present at n Hz)	DP OAE (present/ absent)
1	40	M	36.25	A	500, 1k, 2k	Absent
2	44	M	31.25	A	500, 1k, 2k	Absent
3	33	M	26.25	A	500, 1k, 2k, 4k	Present
4	45	M	32.5	A	500, 1k, 2k	Absent
5	47	M	32.5	A	500, 1k	Absent
6	45	M	23.75	A	500, 1k, 2k, 4k	Present
7	48	M	27.75	A	500, 1k, 2k	Present
8	32	M	26.25	A	500, 1k	Absent
-9	30	M	46.25	A	500, 1k	Absent
10	26	M	33.75	A	500, 1k, 2k	Present
11	27	F	26.25	A	500, 1k, 2k, 4k	Present
12	38	F	26.25	A	500	Absent

Appendix 2

Sl. No.	Consonant Vowel	Sl. No.	Vowel Consonant
1	/pa/	1	/ap/
2	/pi/	2	/ip/
3	/pu/	3	/up/
4	/ta/	4	/at/
5	/ti/	5	/it/
6	/tu/	6	/ut/
7	/ka/	7	/ak/
8	/ki/	8	/ik/
9	/ku/	9	/uk/