PERCEPTION OF TERMINAL INTONATION CONTOUR OF QUESTION-STATEMENT INTONATION IN KANNADA

Register No: 05SLP016

SUMA T.

A Dissertation Submitted in Part Fulfillment of Master's degree
(Speech Language Pathology)
University of Mysore
Mysore.

ALL INDIA INSTITUTE OF SPEECH AND HEARING
MANASAGANGOTHRI
MYSORE-570006

MAY 2007

Dedicated To

My Dearest

Achan & Amma

CERTIFICATE

This is to certify that this dissertation entitled **"Perception of Terminal intonation contour of Question-Statement intonation in Kannada"** is a bonafide work in part fulfillment for the degree of Master of Science (Speech Language Pathology) of the student Registration No. 05SLP016. 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 2007

Dr. Vijayalakshmi Basavaraj
Director
All India Institute of Speech and Hearing
Manasagangothri, Mysore-570006

CERTIFICATE

This is to certify that the dissertation entitled **"Perception of Terminal intonation contour of Question-Statement intonation in Kannada" has** been prepared under my supervision and guidance. It is also certified that this dissertation has not been submitted earlier to any other University for the award of any Diploma or Degree.

Guide

Dr. R. Man jula
Department of Speech Pathology,
All India Institute of Speech & Hearing
Manasagangothri, Mysore- 570006

Mysore
May, 2007

DECLARATION

This is to certify that this master's dissertation entitled **"Perception of Terminal intonation contour of Question-Statement intonation in Kannada"** is the result of my own study and has not been submitted earlier to any other university for that award of any degree or diploma.

Mysore Register No. 05SLP016

May, 2007

ACKNOWLEDGMENTS

I sincerely thank, ***Dr. R. Manjula,*** *Professor, Department of Speech Pathology, AIISH, for her constant guidance and support throughout the dissertation. I thank her for being my constant source of inspiration throughout my student years at AIISH.*

I thank ***Dr. Vijayalakshmi Basavaraj,*** *Director, AIISH, Mysore, for permitting me to carry out this dissertation.*

My ***teachers*** *past and present, you have played a very important role in my life. Each and every one of you has taught me several ideals in life, academic and personal.*

I take this opportunity to thank statistics mam ***(Mrs. Vasantha Lakshmi)*** *who helped me with the statistics work.*

My dearest ***Achan & Amma,*** *your prayers were always with me. It's only through your source of inspiration and love, lam made what lam now. Your hardwork is mirrored now in mine.*

Dearest ***Chechi & Chetan*** *thankyou for providing me moral support whenever I needed it. Chechi, I know I am here just because of you.. Love you lots.*

My sweetest ***Kutoos,*** *you are my constant source of delight, missed you a lot.*

I also would like to thank ***Arun sir*** *and* ***Vandana chechi*** *for helping me, giving me suggestions.*

*Dear **Banu mam,** taking this opportunity to say special thanks to you. I will always remember the care, guidance and encouragement you have given me over the years; it has meant a lot to me.*

*Dearest **Susan,** Thank you for being you, for who you are, however far away and for the words you send to me; I know I can always count on you to cheer me up when I am sad., thank u dear.*

*Dearest **Rahana** (Prince), **Priya** chellam, and **Puru,** what would life be without intimate friends? Thanks to you all I will never have to find out.*

*Dearest **Johnsi** (Johns), my good friend & roommate, I always have a room in my heart for you. I can never forget our days together.*

Dear KD, you are someone on whom I can always depend on. Thank you for all that you have done for me. Your friendship is invaluable to me.

*Dear **Nambi,** I admire you for what you are; you were always ready to lend a helping hand. lam lucky to have you as my friend.*

*Dear **Radheesh** and **Supraja,** how will I forget our fun times, break-time chats, exam preps? The times we have spent together will always remain fresh in my memory.*

*Dear **Kavya** and **Maria,** thank you for helping me in my data collection.*

*Dearest **Sangeetha,** I must acknowledge you, for without your CD player I could never have completed my data collection, Thank u dear.*

I thank Sindushree and Sumitha for helping me prepare stimuli.

***Kamala & Shruthy,** Remembering all the good times I had with u guys, I will cherish them always.*

/ also appreciate the help given so readily by my little friends ***Sneha, Seby, Sara, Deepthi & Rhea,*** *thanks lot dears.*

Brindha & Henta, *we were together for very short period; but its really memorable. Remembering all my beloved seniors (specially* ***Kiruthika mam, Rakhi chechi ,Sneha, Bhuvana, Deema, Rajani, Pooja, JK sir, Sudhakar sir, Hari sir, Ashly)*** *who have been caring and supportive right from my first day at AIISH. You will always be a part of my sweet memories at AIISH.*

Thanks to the ***library staff,*** *AIISH, for the help with the literary treasure hunt.*

Thanks to ***Shivappa*** *&Cofor their kind cooperation throughout.*

I'm *grateful to* ***God*** *for giving me the strength and courage in finishing this study.*

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES & GRAPHS

INTRODUCTION

'A word is not a crystal, transparent and unchanged, it is the skin of a living thought and may vary greatly in color and content according to the circumstances and the time in which it is used.'

- Oliver Wendell Holmes

Suprasegmentals are properties of speech that have a domain which is larger than a single element and includes the features of stress, rhythm, intonation and pause (Lehiste, 1976; Netsell, 1973). 'Intonation' is one of the suprasegmental features which is defined as the variation of speech pitch or fundamental frequency (FO) as a function of time (Collier, 1991). It conveys maximum linguistic contrastivity within the whole range of vocal effect. A speaker can show that he or she is asking for information or confirmation, seeking agreement, or simply making a remark that is indisputable or 'common knowledge', through the intonation of the voice. Even though pitch rarely causes problems with the identification of words, an inappropriate intonation pattern can lead to misunderstanding just as a mispronounced sound can.

Intonation is extensively studied and described for its features and for its variation across languages and attitudes. There is no language in the world which would be regarded as entirely monotonous. In the process of speaking every syllable and every word, the sentence is pronounced with the variation in pitch. Only in very unusual situations we speak with a pitch of the voice which does not change. Various approaches

to the analysis of intonation have been enumerated in the literature. Cruttenden (1997) listed six areas of intonational research which were as follows:

1. The role of syntax in intonation groupings and nucleus placement
2. The set of tones (intonation contours) that constitute the intonational lexicon of an individual language (both in terms of contours and in terms of level sequences)
3. The semantics (abstract meanings) of the set of tones within the intonational lexicon
4. The pragmatics (local meanings) of the set of tones from the intonational lexicon, resulting from " the interaction between abstract meanings of tones and other levels of meaning (lexical, grammatical, gestural)"
5. The realization of rules involved in mapping the tones from the intonational lexicon onto varying stretches of segments which have pre-assigned stresses
6. A comparative study of the preceding areas to refine the intonational typology and knowledge of universals.

Amongst these, intonation in interrogative speech is a topic which has long attracted the researchers. Not only does question intonation vary in different languages but it also varies for different types of questions [Eg: WH, Yes/ No (Y-N) Or Echo questions]. The phenomenon of rise and fall in FO at the end of interrogative sentence is well documented in some languages like French, English, Swedish and Mandarin Chinese (Pike, 1945; Chang, 1958; O'Shaughnessy, 1979; Hadding and Naucler, 1980). The most commonly documented characteristics for questions across languages are high final pitch and overall higher pitch (Hirst and Di Cristo, 1998). The terminal rise in

fundamental frequency is reported to be distinct in Yes/No interrogatives as reported by von Essen (1956) in German, Jones (1957) and Bolinger (1972), in English. For many languages a message constructed of the same segmental material may be perceived either as a statement or a Yes/No question based on the use of intonation (Couper- Kuhlen, 1986).

Need for the study:

In Kannada, which is one of the major Dravidian languages spoken in South India, few studies are available on question intonation. Based on a general observation made by analyzing few interrogatory statements of Dharwad dialect in Kannada language, Patil (1984) reported that the Yes-No questions had a rising pitch at the end. But it is not known whether the final rise in the interrogatives is due to rise in FO over the terminal syllable or a rise in the FO spread over the terminal word in Kannada. Another study by Manjula (1997) found that the FO is higher in the terminal syllable than the terminal word in the Bangalore-Mysore dialect of Kannada. The results also revealed higher percentage of terminal rise in Yes-No interrogatives in the Bangalore-Mysore dialect of Kannada. The study also reported that, the terminal words are significantly longer in duration than the first words and the terminal syllables are longer in duration than the rest of the syllables of the terminal word. However, this study did not investigate the perceptual correlates of terminal syllables or terminal words. The present study attempts to answer the question as to where the loci of FO movement lies on the terminal syllable or on the terminal word, as perceived by normal listeners; in Yes-No interrogatives of Bangalore-

Mysore dialect of Kannada. It also attempts to investigate for the influence of duration, extent and type of rise in FO (gradual/steep/level) in Yes-No interrogatives of Kannada and its perception by normal listeners for its naturalness.

Aims of the study:

The purpose of the study was

a. To determine the loci of FO change (FO over terminal syllable or terminal word) that differentiates question from statement in the Bangalore-Mysore dialect of Kannada using synthesized speech tokens subjected to perceptual judgment.

b. To find the extent of rise (gradual/steep/ level) that would effectively cue statement as question.

c. To find if the duration characteristics (lengthening of terminal syllable or terminal word) would effectively cue a statement from Yes- No question.

The secondary goals of the study were

d. To find the cliticization of the stimuli that would effectively cue a statement as a question.

e. To find the naturalness of the synthetic tokens as perceived by the listeners.

Method:

One model speaker and thirty normal listeners (15 males and 15 females) participated in the study. Model speaker's utterances were subjected to manipulation using PRAAT software. Terminal word and terminal syllable manipulations were done

for pitch (steep rise and shallow rise) and duration (250 ms, 500 ms, 750 ms). The stimuli set thus created were subjected to perceptual judgment by thirty listeners. The task performed by the listeners was to identify the stimuli tokens as questions or statements and to rate the naturalness of the tokens. The responses of the subjects were analysed for:

a) question judgments for the terminal word and terminal syllable manipulated tokens
b) question judgments for pitch manipulated tokens (steep and shallow rise conditions)
c) question judgments for pitch + duration manipulated tokens (250 ms, 500 ms, 750 ms)
d) question judgments for stimuli with different question clitics ('-aa', '-uu', '- ee', '-dri', '-ra\ and '-ri').
e) naturalness ratings for terminal word and terminal syllable manipulated tokens.

Implications of the study:

The results obtained in this study will be useful in

a) Understanding the function of terminal contour in differentiating question from statement specifically in Y-N question context.
b) Application of knowledge in computational linguistic metrics

Limitations of the study:

a) The study needs to be replicated with more number of listeners.

REVIEW OF LITERATURE

Prosody is viewed as declarative ornamentation, functioning to make speech more aesthetically pleasing (Freeman, 1983). Monrad-Krohn (1947) refers to prosody as "melody of language." Prosody is characterized as the use of pitch, loudness and duration to reveal linguistic and affective information during speech (Jakobson & Halle, 1956; Lehiste, 1970). The prosodic devices include pitch patterns (i.e., recurring configurations of consecutive pitch heights) pitch range (i.e., the distance between highest and lowest pitch heights), pitch register (i.e., the general pitch level of a section of speech), stress, volume, tempo, voice quality variation and pause.

Prosody is now used in a number of different ways. Speakers routinely employ prosody to encode linguistic messages based on conventional knowledge about their language system and its sociolinguistic applications (Bolinger, 1978). Unlike emotional attributes, the acoustic correlates of linguistic prosody operate at circumscribed levels of linguistic representation such as the syllable, word, or utterance, directing the manner in which these phenomena have been studied. Prosody is believed to help listeners parse the important information in the continuous stream of speech (Beckman, 1996; Cutler, Dahan, & Donselaar, 1997). In English, prosodic elements are primarily made up of relative alterations of FO and duration (Lehiste, 1970; Cutler et al., 1997). Short-term variations of FO and duration are reported to highlight important phonemes, syllables, and words and long-term relative changes of FO and duration are reported to be used by speakers to inform listeners about phrase and sentence boundaries. Change in FO from the

beginning to the end of a sentence can signal whether a declarative or interrogative sentence is being spoken (Cooper & Sorensen, 1981; Eady & Cooper, 1985; Pell, 2001; Patel, 2002).

Intonation:

Intonation is one of the most endearing aspect of prosodic devices. Intonation is an exclusive term referring to pitch as a function of time or to long term variation in phrases over numerous inflectional shifts (Green, 1964). It is the perception of changes in the fundamental frequency (F0) of vocal cord vibration during speech production (Minifie, 1973). Bolinger (1972) stated that "intonation is the broad undulation of the pitch curve that carries the ripples of accent on its back." Intonation is the rise and fall of pitch of voice, monitored by the laryngeal movements of the speaker.

The functions of intonation are traditionally divided into two groups: "attitudinal" and "linguistic" (Crystal, 1969; Lehiste, 1970). More recently, these have also been referred to as the "linguistic" and "paralinguistic" dimensions of intonation (Tench, 1996). Intonation allows the speaker to convey his emotions. The intonation contours that reflect attitudes are conventional and hence they may differ from language to language in form or meaning. In terms of its grammatical functions, intonation provides a sentence with a variety of meanings depending on the emphasis placed by the speaker. In addition to the all- grammatical and all-attitudinal views on intonational functions there is a third way of asserting homogeneity of these functions. This approach was

started at Brazil in the 1970s and its main claim is that intonation is not a property of the sentence or the sentence- based utterance but of discourse and so its functions are neither grammatical nor attitudinal but "discoursal" showing informational status and regulating conversational behaviour (Brazil, Coulthart and Johns 1980; Brazil, 1995).

All spoken languages have inherent patterns of intonation. Intonation studies have been neglected in many languages. This neglect was noticed by Labo (1961) who said that "inspite of its importance in many languages, intonation signals have been the last to be analysed linguistically, and still remain controversial in the few languages in which it has been analysed.

Intonation has been studied from different perspectives. Among the studies of intonation some describe the units of intonation of a language and some describe the intonation contours along with their attitudinal meanings. There are a number of theories about how the fundamental components of intonation should be described either as: a) discrete elements or variable ones b) wholes (tonal contours) or parts (tonal sequences).

In the analysis of intonation, some investigators use the 'contour analysis' (Eg: the British investigators) and some other investigators use the levels analyis (Eg: the American investigators). In level analysis, levels of stress or prominence, pitch declination throughout the intonation group, and pitch at terminal junctures (such as falling, rising, or level pitch) are considered. In contour analysis, the location of 'nucleus' and the pitch movement or contour movement with reference to the nucleus are

considered. Different types of intonation contours are identified with reference to the 'nucleus' of an utterance. In British English, the intonation is usually described in terms of contour shape using a fall/rise opposition. The fall rise behaviour of FO is also studied extensively in American English. This is described in many studies on the perception of synthetic speech and the phonetic structure of intonation [Studdert-Kennedy & Hadding (1973), O'Shaughnessy (1979) and Ladd (1983)]

In contrast to the traditional form-based sentence-level approaches, recent research in the United States and elsewhere has been taking a more functional, discourse-based approach, focussing on the role of intonation in natural communication and interaction based on the intonation unit (IU) as the basic linguistic unit (Chafe, 1993; Dubois, Cumming, Schuetze-Coburn, and Paolino, 1992, 1993; Couper-Kuhlen & Selting, 1996 b). In addition to the current interest in function-based approaches, another major focus in recent years has been the attempt to integrate the perceptual properties of intonation using instrumental or acoustic approaches.

Fundamental frequency contours in speech are determined by a number of different factors involved during speech production. Two of the most commonly studied factors are intonation (suprasegmental) and segmental structures. However, the interaction between segmental and suprasegmental structures have often been less explored and less understood (Silverman, 1986). At the acoustic level, FO contours are studied in terms of direction of movement within the utterance and the behaviour of the terminal portion of the utterance, where as some others have considered the variation of

FO across 'intonation groups' (Cruttenden, 1986) and phrase boundaries (t'Hart and Cohen, 1973). Apart from pitch, duration (length) and loudness are also considered as important attributes of intonation (Cruttenden, 1986). Some theories and approaches concerned with intonation have focused on the semantic correlates of statements and questions (O'Shaughnessy, 1979).

Duration is also considered as one of the dimensions of intonation. It is of significance at various segment levels, viz., syllable, word, utterance and pause and it is of importance as it exerts certain constraints over rhythm and tempo which indirectly influence intonation. The total duration of utterance including the duration of pause constitutes the total duration of the intonation utterance.

Loudness, like the feature of pitch and pitch range, is observable throughout the utterance as an indexical feature. However, it is also restricted to the nuclear word where the prominence or the strength of the meaning of an utterance is located. Often the degree of loudness observed on the nuclear word is considered important in interpreting the intonation pattern of a sentence. In all other places it is dependent on pitch. That is, it is correlated with the change in pitch of the syllables. It is one of the features responsible for accentuation, the other features being pitch and duration.

Intonation: Production aspects

The study of intonation is complicated due to the fact that intonation can be studied from various perspectives and using various methods. Several experiments have been conducted to understand the acoustic characteristics of intonation.

'Intonation' is viewed as the product of a combination of different prosodic systems of pitch contrasts. The perceived rise and fall of pitch over linguistic unit is called intonation (Freeman, 1983). Pike (1945) emphasized the role of fundamental frequency or pitch in intonation: "Every sentence, every word and even every syllable is given some pitch when it is spoken." "Even a sound is produced at certain pitch level." However, according to Denes (1959), it is possible that variations of other acoustic characteristics like duration, intensity also contribute to intonation. Researchers have been attempting to specify the fundamental frequency contours of speech based on linguistic specifications. However, this is rendered difficult owing to the contribution of many linguistic features to the FO contours, and the trading relation of FO with other correlates such as duration and intensity (O'Shaughnessy and Allen, 1983)

"Stress" is defined as the perception of some linguistic units as emphasized or prominent in contrast to surrounding units (Freeman, 1983). Armstrong and Ward (1926) state that stress and intonation are so closely connected that it is often difficult to decide whether stress or intonation or the combination of the two is responsible for certain

effects. The role of pitch in stress is difficult to determine because of its association with intonation. But it has been pointed out that in general there is a tendency for high relative pitch to carry an implication of greater stress (Fry, 1958). "Any study of the meaning of pitch contours requires first some decision on the nature of stress, for the very taxonomy of intonation depends on how independent pitch and prominence are considered to be (Ladd, 1983)." In the British tradition, sentence stress is commonly called the "nucleus". In British English intonation, the nucleus is described over a range of levels. They include auditory ('prominence' by Halliday, 1967), acoustic ('largest pitch movement' by Kingdon, 1958), rhythmic ('stressed syllable' by Schmerling, 1976), functional ('focusing' by Chomsky, 1970). All these, irrespective of the subtle differences in their proposed definition indicates that there is a region of intonation contour that focuses attention on a stretch of utterance (ranging from syllable to breath group).

Lieberman (1959) reported that stressed vowels in English had a higher F0 72% of the time, higher amplitude 90% of the time and longer duration 70% of the time when compared to unstressed vowels across words. Higher F0 and amplitude appeared to be stronger 'cues' than duration, in that no stressed syllable had both a lowered F0 and decreased amplitude. In addition to this, he found a 'trading effect' where, in cases of lowered F0, there was an increase in amplitude, and vice-versa. He concluded that increased F0 was a stronger 'cue' than amplitude. Lieberman's results disagree slightly with that of Fry (1955; 1958) which showed that both duration and amplitude affected stress judgements of synthetic speech, and that change in duration, but not amplitude was enough to alter judgements. Fry (1958) also reports that the degree of difference in F0

between stressed and unstressed syllables in the same word had no effect, as long as the difference was discernable. Klatt (1976) reported that the average duration of a stressed vowel was approximately 60 ms longer than an unstressed vowel, and that phrasal position also affected intrinsic duration. Thus the various ingredients of intonation like the pitch, duration and amplitude play a salient role in bringing a syllable or utterance into perceptual prominence.

Thorsen (1978) investigated the acoustic attributes of Danish intonation. The relationship between linguistic stress and fundamental frequency and the intonation contours in short declarative sentences, interrogative sentences and non-terminal pauses in Advanced Standard Copenhagen Danish were established as the basis of acoustical analysis of speech of four subjects. The intonation contours in this language varied with sentence type from smoothly falling (statements) to level (intonation questions). The stress group patterns were invariant, a low stressed syllable followed by a tail of high falling unstressed syllables, but the absolute distance between the stress and the first post-tonic syllable decreased progressively from the first to the last stress group in the sentence.

Other prosodic systems comprise independently varying vocal effects based on combinations of pitch, duration and loudness parameters in specific ways (e.g. rhythmicality) or on contrast in silence (the system of silence pause) (Crystal, 1969).

Bolinger (1972) states that there are three features of intonation which have similar functions in all languages. They are

1. *Range:* The range conveys emotions. When one is excited the voice pitch extends upwards. When one is depressed, he/she speaks almost in a monotone. As this feature is not usually under voluntary control, it is considered another instance of instinctive gesture.
2. *Direction:* It is usually connected with a pause. The two together are the punctuation marks of speech. The tendency in all languages, in making statements, is to have a fairly high pitch towards the beginning of a sentence and then to drift down to the lowest pitch at the end. In questions that are answered by "Yes" or "No", the direction often tends to be up all the way.
3. *Relative height:* It is associated with the importance given to a particular word or words in a sentence.

Pike (1945) divides intonation characteristics roughly into two types:

1. The contours which are completely colorless in meaning. This is known as intentional minimum of speech which serves as a mould into which all minimum sentences may be poured so that the intended utterance is achieved.
2. The other intonation characteristics are affected or caused by the individual's physiological and psychological state like anger, happiness, excitement, age, sex, etc. These are the ones which help one to identify people and to ascertain how they are feeling. Whenever a certain sequence of relative pitch is heard, one may

conclude that the speaker means certain things. One sentence may have several contours and a single contour may have several meaningful parts.

Most research studies have focused on identification of intonation by exclusively investigating the *production* domain. However there are indications that the 'visible' manifestation of pitch does not automatically correspond to the way it is perceived by the listener. Thus, some apparently major pitch events may play a negligible role in perception while, conversely, a seemingly minor phonetic detail may prove indispensable (Cruttenden, 1992).

Intonation: Perception aspects

In dealing with the perceptual aspects of intonation from a phonetic or acoustic perspective, two main aspects need to be considered: first, the human ability to perceive the physical properties of frequency, duration and intensity and second, the psychological response to various acoustic stimuli.

Investigators have reported that human listeners, starting from birth, attend and react to speech sounds differently than to other acoustic signals. It is reported that the human brain has a number of devices or property detectors that respond selectively to particular types of acoustic signals and perceive speech signals as a mode different from other sounds (Chun, 2002). Although psychophysical investigations have been able to isolate the individual elements of fundamental frequency, duration and amplitude to determine listeners' ability to perceive them, psychoacoustic research on the

suprasegmentals properties of pitch, length and loudness has suggested that listeners use a combination of these components in order to make optimal perceptual judgments (Chun, 2002). Cruttenden (1997) found that in English, the three features (pitch, duration and loudness) are of importance in bringing syllables into prominence, pitch being the most efficacious and loudness the least so.

Studies on perception of intonation contours suggest that fundamental frequency or pitch plays an important role in how listeners attribute speaker's emotions on the basis of different acoustic cues (Ohala, 1984; Ladd, 1985). Sweet (1982) equated intonation with pitch or fundamental frequency of the voice and stress with loudness. He described the types of intonation as level, rising or falling and the level tone. The level tone may be either high or low and the other tones may begin in a high or low pitch. Garding and Abramson (1965) studied listeners' abilities to perceive and discriminate different intonational forms using synthesized speech. They attempted to investigate the intonation contours in American English that could be identified by native speakers on the basis of fundamental frequency alone. They tried to determine the extent to which the pitch-curve of a given contour could be synthetically changed before being perceived as different. The FO movements of five intonation contours were varied in a systematic fashion and subjects were required to decide whether pairs of contours were "the same" or "different". The results showed that each contour had a considerable margin within which changes could be made without any effect on perception as long as these changes did not disturb a certain pattern.

The specific focus of several studies has been on the parameters of intonation contours. In general, listeners identify a contour type by attending to certain aspects of the fundamental frequency (FO) pattern. Several studies were carried out to identify parameters of FO patterns by Hadding-Koch and Studdert-Kennedy (1964). They presented American- Swedish listeners with a series of synthetic stimuli. The subjects were asked to identify each stimulus as statement or question. The results indicated that both the group of listeners identified the contours by attending to three cues: the fundamental frequency of the "peak" and the "turning point" in FO, and the extent of the terminal glide, also suggesting a trading relationship among these three cues. On similar lines, t'Hart, Collier and Cohen (1990) studied Dutch intonation and reported that perceptually at least ten pitch movements are distinguishable in Dutch language. These distinctions were based on (a) whether the movement was a rise or a fall, (b) whether it extended over one or more than one syllable, (c) its temporal position in the syllable, and (d) its excursion size. The duration of sounds, syllables, words or phrases will have their share in the prosodic system of a language

Intonational phonology suggests regular features in intonational structure, both in terms of the scaling of tonal targets as well as concerning their position. Maeda (1976), for instance, reported that speakers vary little in the low pitch at the end of utterances. Liberman & Pierrehumbert (1984) showed that the height of accents is highly predictable from the context they appear in. Increasing the pitch height of an accent has been reported to increase its perceived prominence (Rietveld & Gussenhoven, 1985; Terken, 1991; Kohler & Gartenberg, 1991) and emphasis (Nolan, 1995; Ladd & Morton, 1997).

Eady and Cooper (1986) conducted a study to determine the manner in which the location of linguistic focus influences intonational attributes of duration and fundamental voice frequency (FO) in matched statements and questions. Speakers orally read sentences that were preceded by aurally presented stimuli designed to elicit either no focus or focus on the first or last noun phrase of the target sentences. They reported that the FO topline was dependent on sentence type and focus location. For sentences with neutral or sentence-final focus, the difference in the FO topline between questions and statements was evident only on the last key word, where the FO peak of questions was considerably higher than that of statements. For sentences with focus on the first key word, there was no difference in peak FO on the focused item itself, but the FO toplines of questions and statements diverged quite dramatically following the initial word. The statement contour dropped to a low FO value for the remaining part of the sentence, whereas the question remained quite high in FO for all the subsequent words. In addition, the FO contour on the focused word was rising in questions and falling in statements, regardless of focus location. Computer-aided acoustical analysis of word durations showed a localized, large magnitude increase in the duration of the focused word for both statements and questions.

Intonation in interrogatives:

Intonation in interrogative speech has been a topic of interest since many years. Austin (1975) introduced the notion that human speech can be conceived as consisting of numerous "speech acts". Even when they contain the same words (i.e. when they have

the same prepositional meaning), statements and questions differ as speech acts, since a question involves a certain type of intonation (Searle, 1969). The illocutionary act of "posing a question" maybe signaled lexically by the use of WH words; by a change in word order; or prosodically through the use of intonation (Couper-Kuhlen, 1986).

Studies on the perception of intonation patterns has shown that questions and statements can indeed be distinguished from each other on the basis of a rising or falling terminal FO contour alone (Uldall, 1962; Majewski and Blasdeil, 1969). In a study of the perception of sentence intonation in Danish, Thorsen (1980) found that the fundamental frequency (pitch) was closely correlated with subject's judgments of utterances as declarative or interrogative. The most steeply falling intonation contours were identified as being declaratives; the least falling ones were identified as interrogatives. Other studies indicate that listeners also rely on other aspects of the FO contour when differentiating questions and statements (Hadding-Koch and Studdert-Kennedy, 1964; Studdert-Kennedy and Hadding, 1973, 1974).

Several researchers have examined the characteristics of statement-question prosody that are functional in perception. Majewski and Blasdeil (1969) recorded the word *farmer* spoken as a statement or as a question. The word was then synthesized and presented to participants for identification. Their conclusion was that questions and statements could potentially be distinguished from each other based on terminal FO contour alone. Studdert-Kennedy and Hadding (1973) suggested, however, that listeners also rely on other aspects of the FO contour in differentiating statements and questions. In

their study, they imposed synthetic contours on naturally spoken utterances and manipulated the FO of the contour. The perceptual judgments of statement or question indicated that listeners were influenced not just by the terminal FO but the entire FO contour.

Questions are often said to be distinguished from statements by a terminal rise in fundamental frequency as against a terminal fall. Questions may also be distinguished by a comparatively high FO through out the utterance (Hadding-Koch and Studdert Kennedy, 1964). An early work (Studdert-Kennedy & Hadding, 1973), which sought to illustrate the importance of incorporating intonation in perceptual judgments, compared intonation in speech and nonspeech acts. The word *November* was resynthesized to create 72 intonation contours that varied in FO at peak stress and word end. In addition, analogous sine waves and frequency-modulated pulses of varying FO were constructed. Listeners were asked to judge whether the stimuli were representative of a statement or a question. It was concluded that peak FO and perceived terminal glide were most utilized by listeners in making judgments as to whether an utterance was a statement or a question.

Lieberman (1967) had participants read statements like *Joe ate the soup* and echoic questions like *Joe ate the soup?* The echoic question had the same word order as the statement, but it was interrogative in nature, sometimes marked by an element of surprise. The recordings were analyzed for pitch information. The statements were characterized by a falling terminal fundamental frequency (FO), whereas the questions were marked by a final rise in terminal FO.

Danes (1960) noted that in Czech language a special interrogatory contour is used only in Yes or No questions which did not otherwise differ significantly, either lexically or grammatically, from statements. Hadding-Koch (1961) stated that in Spanish, intonation was the only feature that distinguished an interrogative question from a statement. Both statement and question had the same word order. They differed only in that the question concluded with a rising pitch while statements concluded with a falling pitch.

O'Shaughnessy (1979) analyzed sentences produced by four speakers and identified three critical regions: the initial, medial, and final accented syllables of questions. He observed that all three syllable positions were characterized by a rising F0 contour. He concluded that the question intonation affects the F0 contour of the entire sentence, and is not limited to a rising contour at the end of the utterance. An interaction between speaking rate and intonation in French by Canadian speakers was reported by Ryalls, Dorze, Lever, Oullet, & Larfeuil (1994). They looked at the duration of matched statements ("The bird is singing.") and echoic questions ("The bird is singing?"). Although questions were spoken at a faster rate than statements, the final syllable durations were significantly longer for questions than statements. This interaction between speaker rate and sentence type has not been corroborated in English.

The focus of research into question/statement differentiation has been mainly on production factors (e.g. Cooper and Sorenson, 1981; Cutler and Ladd, 1983; Lieberman

& Pierrehumbert, 1984). Early descriptions of intonation in English state that FO (or voice pitch) has a generally rising pattern in Yes-No questions, in contrast to the generally falling pattern of statements (Pike, 1945). The shape of the FO contour at the end of a sentence is said to be a very important determining factor. Lieberman (1967), for example, claims that the FO contour during the final 150-200 ms of phonation can reliably differentiate questions from statements (or, in his terms, marked and unmarked breath group).

There are few studies which report on the perceptual aspects of question intonation. Makarova (2001) conducted a perceptual experiment in which Japanese and Russian subjects were asked to identify two-syllabled re-synthesized stimuli with modified rising-falling contours as exclamations, interrogatives or declaratives. The results revealed that the increase in the height or magnitude of a rise led to a decrease in the perception of stimuli as declarative by both groups of listeners.

Another study by Rathcke (2006) in Russian also investigated the perceptual aspects of question statement in Russian. It was observed that the major perceptual cues for this category distinction were the FO-peak alignment and the slope of the rise. The primary perceptual cues for questions were a steep rise and a late peak alignment at the offset of the accented vowel, whereas the more gradual rise and early FO-peak alignment at the onset of the accented vowel were strong cues for a declarative mode. They also reported that the height of the FO-peak had no influence on the category distinction.

The questions or interrogatives in many languages are generally classified into four types (Schiffman, 1979; Shankarabhatta, 1978). They are:

- **WH questions:** Here, the information is sought from the listener using different question words like What, How, Where, Why, Who, Which, Whom.
- **Yes-No questions:** questions belonging to this category expect a Yes or No answer from the listener.
- **Alternate questions:** here, the listener has to choose from two alternate answers offered by the questioner through his question.
- **Echo questions:** questions which try to get confirmation from the listener by added structures such as "Isn't it?" "That's all" or in other words questions

where a questioner echoes his own utterance to confirm an answer which he expects.

Among these types of questions, a Yes-No question has a distinct rising terminal contour. Commenting on the types of Yes-No questions in English, Lieberman (1967), stated that there were 'simple' Yes- No questions in English. These were 'normal' Yes-No questions said neutrally where only information was sought. A second type referred to as 'intonation questions' was a statement with final marked breath groups and rising fundamental frequency. These were considered as ambiguous and could be interpreted as paraphrases of metalinguistic sentences (echo questions). Semantically these intonation questions either indicated surprise or doubt. The formation of questions across languages varies. In Japanese a question that takes the answer 'Yes' or 'No', can be formed in two ways. Abe (1955) reported that a statement like the expression *Darekakita* (Somebody came) can be turned into a question by the addition of the interrogatory particle *Ka.* The

sentence becomes *Darekakitaka?* However, *Darekakita* can also be turned into a question by concluding the sentence with a rising intonation. A similar effect is also reported to take place in Chinese by Chang (1958), who studied the Chengtu dialect and noted that a particle with a rising intonation at the end of the sentence indicates that it is a question. The particle has a falling intonation when the sentence is a statement.

A typical pattern of rising contour distinguishes Yes-No questions from statements where a falling FO contour is seen (Magdics, 1963; Micky; 1977; Manjula, 1997). Yes-No questions are also reported to end in other ways, not uncommonly in an intonation fall. (Bolinger, 1978; Lee, 1980). Thus the terminal contour of question intonation has a significant role. However the location or beginning of a terminal contour in a given segment and its potential to cue or differentiate a sentence from question is largely unexplored.

Vaissere (1983) observed that the terminal contour of the last syllable is an important feature in distinguishing simple declarative statements from Yes-No questions in a number of languages. This view was confirmed by Contini and Boe (1975) for French, Magno Caldognetto (1978) for Italian, Nishinuma (1979) for Japanese, Dascalu (1979) for Romanian and Thorsen (1980) for Danish. Other than the pitch movement, duration characteristics and intensity changes in the final words of interrogatives are also listed as being important in differentiating statements and questions by various investigators (Abe, 1955; Lieberman, 1967; Uldall, 1964). The tendency to lengthen the final elements in an utterance is found to occur in many languages as reported by

different investigators. In French, English, German and Spanish (Delattre, 1966) in Swedish (Lindblom, 1968) and in English (Umeda and Quinn, 1981). Cruttenden (1986) reported that in some East European languages (Russian, Czech, Hungarian and Romanian) the increase in pitch was observed on the prominent syllable of Yes-No questions which was not necessarily located in the terminal contours of languages. Gosy and Terken (1994) stated that in Hungarian, the distinction between statements and Yes-No questions was not expressed by lexical or syntactic means, but only by prosody. Yes-No questions showed a pitch change in the penultimate syllable, regardless of whether or not the syllable was stressed. They further observed that the timing of the pitch peak in the penultimate syllable, the peak height, the shape of the pitch change and the presence of an additional pitch peak in the beginning of the utterance affected its interpretation as a statement or a question.

Intonation in Kannada:

There is dearth of studies in Indian languages addressing the location and direction of FO change in terminal portion of a speech segment. Kannada is one of the major languages of Dravidian family, especially south Dravidian family (Hiremath, 1961). There are several dialects, however in the present study the Bangalore Mysore dialect of Kannada has been considered.

Manjula (1979) studied intonation in Kannada under nine emotional conditions using 36 sentences. She had concluded that "The sentences in Kannada with emotions are expressed with a final fall in the intonation pattern." A sentence with a specific emotion

can be expressed with more than one type of intonation pattern and a single intonation pattern may be used to express sentences with different emotions. Nandini (1985) studied the kind of intonation patterns that are used by the speakers of Kannada language in expressing various emotions. She found that neutral, jealousy, hesitation, request are the types of sentences that have shown a final fall (either a gradual or steep) and other sentence (question, anger + question, frustration, accusing type have shown a final rise (gradual/steep patterns). Investigations of Rathna et al. (1981) showed that a relative increase in intensity, steepness of intensity rise, a pause before the word and a large duration of the word are the features contributing to stress in Kannada language.

Nataraja (1981) compared the intonation contour in four Indian Languages (Kannada, Tamil, Gujarathi and Hindi) under five emotional conditions a) Anger b) Joy c) Jealousy d) Neutral e) Mercy. He concluded that same intonation contour may be used to express different emotional conditions and further, the same patterns, or contours are same across the languages used. He also concluded that there seems to be common or 'universal' intonation contours across languages studied.

Manjula (1997) studied the features of intonation and stress in "Question word (WH)' and "Yes-No" interrogatives in the spoken form of Bangalore-Mysore dialect of Kannada language. In this study, the fundamental frequency, duration and intensity parameters over terminal word and terminal syllable of Yes-No questions were studied in detail. A higher percentage of occurrence was reported for terminal rise in Yes-No questions. The FO was also reported to be higher in the terminal syllable than the terminal

word. The terminal word was longer than the initial words and the terminal syllables were longer in duration than the rest of the syllables of the terminal word. This study was from a production perspective.

Yes-No questions in Kannada:

In Kannada, usually the long vowels 'ee', 'aa', 'oo', are added to the end of the Yes-No interrogatory sentences (Shankara Bhatta, 1978). "Any sentence can be made interrogative by adding the Yes- No question marker 'aa' to almost any constituent usually the finite verb"(Schiffman, 1979). Sridhara (1990) made a distinction between *neutral* Yes-No questions, which do not expect a particular type of answer and *leading* Yes-No questions, where an affirmative answer is expected.

Neutral Yes- No questions:

These are formed by the clitic '-aa' (dialectally '-ee' or '-oo') attached to the end of the statement. There is no change in the word order, but cliticization of the morpheme causes the final vowel to drop. Vocative clitic changes to indicate the sex of the addressee and the nature of relationship between the speaker and the addressee (Sridhara, 1990).

Eg. Vocative clitic '-oo' is used when the addressee is female and the relationship is casual, '-ri' or '-ra' is used when the addressee is a male or female and the relationship is polite.

Leading Yes-No questions:

Positive leading questions (that is, questions expecting an affirmative answer) are formed by using the particles alia (variants: allava:, allave:, allavo: all formed with question clitic a: or e: and very rarely o:)

There are very few studies of intonation in Kannada which have analysed the acoustic aspects of intonation. There is no literature on the perceptual aspects of intonation in Kannada language. The present study was an attempt to analyze the terminal contour in Kannada language, which is one of the major Dravidian languages spoken in southern state of India. This study attempts to find and determine the loci of FO change (FO over terminal syllable or terminal word) which has the potential to differentiate a question from a statement and also the extent of rise (steep/shallow/level) that would effectively cue a statement as question, in the standard dialect of Kannada, using synthesized tokens subjected to perceptual judgment.

METHOD

The present study aimed to find the loci of FO change (FO over terminal syllable or terminal word) which served as cues to differentiate a Yes-No question from a statement. It also investigated the extent of rise (steep/shallow/level) in FO that would effectively cue a statement as Yes-No question. The study was conducted in Bangalore-Mysore dialect of Kannada language, using synthesized tokens which were later subjected to perceptual judgment.

The study was carried out in the following stages:

1. Subject selection
2. Preparation of stimuli
3. Perceptual evaluation

1. Subject selection:

Subjects in the study included

a) One model speaker and

b) Thirty normal hearing naive listeners

Criteria for selection of model speaker;

One adult female speaker of age 21 years, who was a qualified speech-language pathologist and a native speaker of Bangalore-Mysore dialect of Kannada, with a working knowledge in the area of prosody, was selected. The speaker was screened for hearing, speech-language problems and neurological deficits.

Criteria for selection of listeners:

Thirty listeners (15 males and 15 females) in the age range of 20-40years, whose hearing sensitivity was within normal limits and who did not have any history of any loss of hearing or ear pathologies, were selected. All subjects were native speakers of Bangalore-Mysore dialect of Kannada.

2. Preparation of stimuli:

Step 1:

Twenty sentences, in the Bangalore-Mysore dialect of Kannada, which could be expressed either as a Yes-No question or as a statement, were selected. The mean length of these utterances varied from three to six words per sentence. The sentences were selected in consultation with a linguist who was well versed in Kannada language. The sentences when added with the clitic marker (clitiization) '-aa', '-uu', '-ee', '-ra', '-ri', and '-dri' could be expressed as a Yes-No question. The sentences with addition of clitic markers were spoken by a 'good' speaker as a statement. The terminal portion of these statements were manipulated to alter the pitch and duration synthetically in order to test for the extent of manipulation that would be required for the statement to be perceived as a Yes-No question by a listener. The sentence stimuli are given in Appendix A.

Step 2:

A 'good' speaker was selected as a model speaker. The 'good' speaker in this study was a qualified speech-language pathologist with a working knowledge in the area

of prosody. The speaker was instructed to utter twenty sentences as declarative statements. Trials were given until the speaker and the investigator was satisfied that she had uttered the stimuli as declarative statement. The best uttered statement was retained for final analysis. The speech samples were recorded using a stereo tape recorder with unidirectional mic, which was kept at a distance of less than eight inches from the mouth.

Step 3:

In order to check whether the speaker had effectively expressed the sentence as declarative statements, the same was subjected to perceptual judgment by three judges. The judges were speech-language pathologists with working knowledge in the field for not less than five years. The judges were asked to listen and indicate on a checklist provided by the investigator whether the utterance was a statement or Yes- No question. Twelve sentences that were judged as being most effectively produced as declarative statements by all three judges were retained for the final analysis by the investigator. These twelve sentences were grouped into sets of six with two statements in each, which when manipulated could be categorized as questions with cliticization of '-aa\ '-uu', '-ee', '-ra', and '-ri' and '-dri'.

Step 4:

The twelve original statement tokens spoken by the speaker were selected and manipulated through a multistage procedure using PRAAT speech analysis software package (Boersma & Weenik, 2000).

(a) First the beginning and end of the 'terminal syllable' in each statement was marked in the waveform. The FO of the final syllable of the twelve statements was manipulated in three ways:

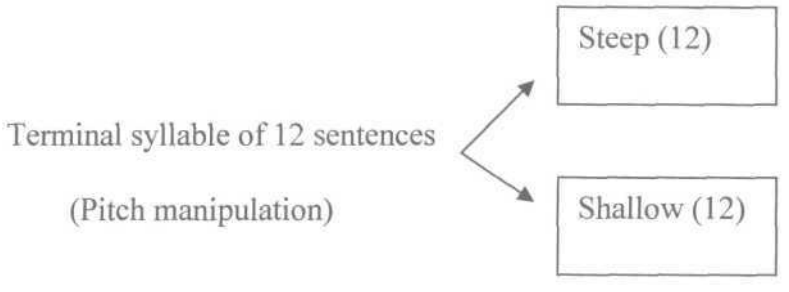
a) The FO of terminal syllable was manipulated to a steep rise.

b) The FO of terminal syllable was manipulated to a shallow rise

c) The FO of terminal syllable was kept level without any change in FO.

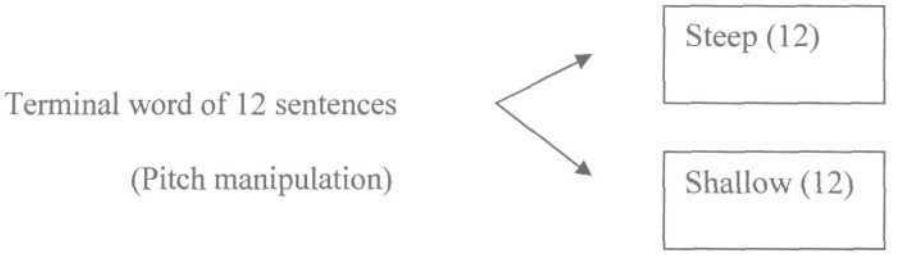
The 'steep' rise is operationally defined in this study as a rise of greater than 100 Hz from syllable initial to syllable final position of the terminal syllable. The 'shallow' rise is defined as a rise of 50-100 Hz from syllable initial to syllable final position of the terminal syllable (Manjula, 1997).

The pitch synchronous overlap and add (PSOLA) technique was used to achieve F0 shifts without affecting the tempo of the recording (Mouliner and Charpentier, 1990).This process resulted in forty eight synthesized tokens.



(b) Next, the beginning and end of the 'terminal word' was marked in the waveform and also in the sample and was manipulated similarly in 3 ways. In this study the 'steep' rise for terminal word was considered as a rise of greater than 100Hz from

initial syllable of the terminal word to the final syllable of the terminal word. The 'shallow' rise was considered as a rise of 50-100 Hz from initial syllable of the terminal word to the final syllable of the terminal word. This process resulted in another forty eight synthesized samples.

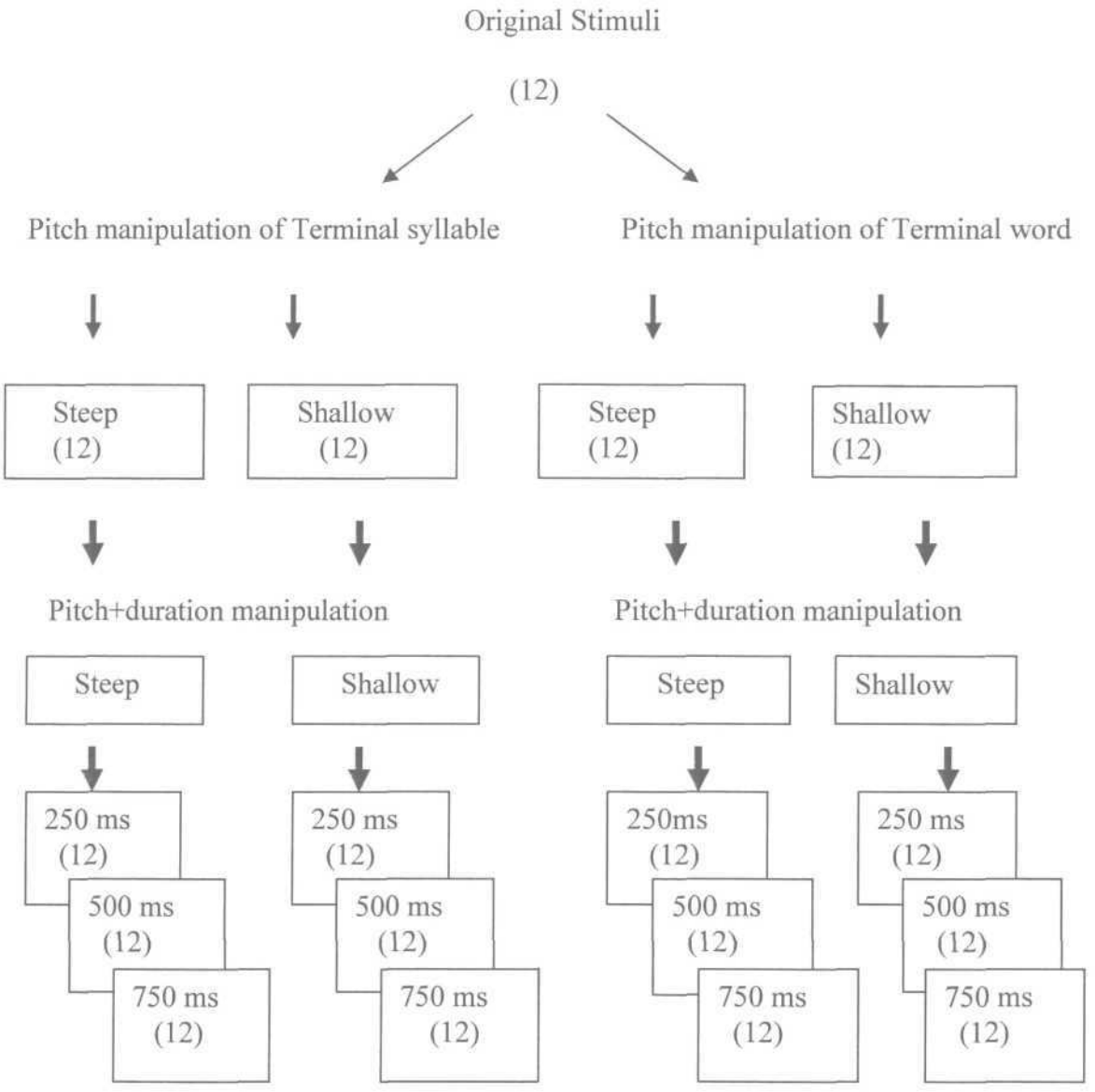


Thus 48 manipulated tokens (24 for terminal syllable + 24 for terminal word) constituted one stimuli set.

Another set of stimuli consisting of pitch+duration manipulated tokens were created using the same PRAAT software. Here each of the pitch manipulated (both terminal syllable and terminal word manipulated tokens were lengthened by 250 ms, 500 ms, 750 ms resulting in a total of 144 manipulated tokens. These 144 manipulated tokens, together with 48 pitch only manipulated tokens and 12 original un manipulated stimuli tokens, constituted total stimuli of 204 tokens.

Conditions		**No. of Stimuli**
Original un manipulated		12
Manipulated	Pitch	48
	Pitch+duration	144
Total		204

Fig 1: Overview of manipulated tokens



The Listening Task:

The 204 tokens (192 manipulated+ 12 un manipulated) were randomized and transferred to a compact disc for presentation to listeners. Perceptual identification of the samples was carried out by individual listeners in soundproof rooms and the samples were presented through headsets. Listeners were asked to identify each of the tokens that were presented through earphones one by one, as statements or questions and also rate how natural they felt the tokens were (unnatural/ fairly natural/ natural) on a response sheet given to them. The sample of response sheet is given as Appendix B. When listeners were unsure, the samples were played twice and they were asked to make their best possible judgment.

Analysis:

The responses identified as questions by each of the listeners were calculated and raw scores were analysed using appropriate statistical measures. Repeated measure ANOVA was used to find the difference in question responses of listeners for various manipulation conditions and Paired sample t test was used to see the pair-wise significance between different manipulation conditions.

RESULTS AND DISCUSSION

Review of literature shows that there are several languages where a message constructed of the same segment of speech may be perceived either as a statement or a question solely based on the function of the terminal intonation. There is dearth of studies in Indian languages addressing the location and direction of FO change in terminal portion of a speech segment. Hence, this study was taken up to examine and analyse aspects related to the terminal contour in the speech of a Kannada speaker using Bangalore-Mysore dialect.

The objectives of the study were:

a) To determine the loci of FO change (FO over terminal syllable or terminal word) which helped to differentiate Yes-No question from statement in the Bangalore-Mysore dialect of Kannada using synthesized speech tokens which were prepared for the study and subjected to perceptual judgment.

b) To determine the extent of rise (gradual/steep/level) that would effectively cue statement as Yes-No question.

c) To determine the contribution of duration characteristics (lengthening of terminal syllable or terminal word) in effectively cuing and differentiating a statement from Yes- No question.

As a secondary goal the synthetic tokens were perceptually analysed by the listener for their 'naturalness'.

To achieve these objectives, the sentence stimuli were selected such that when the terminal portion were altered in terms of pitch and duration, it could be expressed as Yes-No questions with addition of clitics like '-aa', '-uu', '-ee', '-ra', '-ri' and '-dri'. Multistage manipulations were done to alter the stimuli as per the requirements using PRAAT software. The terminal words and syllables of sentences were initially manipulated for extent of rise in pitch (steep vs. shallow) and later pitch+duration manipulations (250 ms, 500 ms and 750 ms) were carried out. A total of 204 synthesized tokens thus prepared were given to thirty listeners for perceptual evaluation, and were asked to make a forced binary choice judgement as to whether they perceived the tokens as a statement or question. They were also asked to rate the naturalness of the synthesized tokens. The responses to the token by each of the listener for each of the manipulation condition was calculated which formed the raw data. The results are drawn based on how the listeners perceived the synthesized tokens.

The results of the study are presented under the following sections:

1. Perceptual loci of F0 {Terminal Word (TW) vs. Terminal Syllable (TS) manipulated tokens}
2. Perception of extent of rise {Steep (ST) vs. Shallow (SH) rise} of manipulated token.
3. Perception of pitch+duration (3 durational increments)
4. Perception of 'naturalness' of synthesized tokens.
5. Perception of resynthesized tokens with 6 different clitic markers.

Table 1: Mean and Standard deviation of responses across different sets of manipulation conditions.

S.NO.	CONDITIONS	MEAN	STANDARD DEVIATION
1	TWSTP TOTAL	6.1000	1.8448
2	TWSTP+*250 TOTAL	7.2333	2.2079
3	TWSTP+*500 TOTAL	5.3667	1.9205
4	TWSTP+*750 TOTAL	6.1333	2.0297
5	TWSHP TOTAL	5.3667	1.3515
6	TWSHP+*250 TOTAL	5.5000	2.1456
7	TWSHP+* 500 TOTAL	5.4000	1.4527
8	TWSHP+*750 TOTAL	6.0000	1.6815
9	TSSTP TOTAL	5.0333	1.6709
10	TSSTP+*250 TOTAL	6.4333	2.3879
11	TSSTP+*500 TOTAL	4.3333	1.7486
12	TSSTP+*750 TOTAL	5.1667	1.5332
13	TSSHP TOTAL	4.1667	1.6206
14	TSSHP+*250 TOTAL	4.8000	1.7695
15	TSSHP+* 500 TOTAL	4.5667	1.9945
16	TSSHP+*750 TOTAL	4.5667	1.7943

Note:

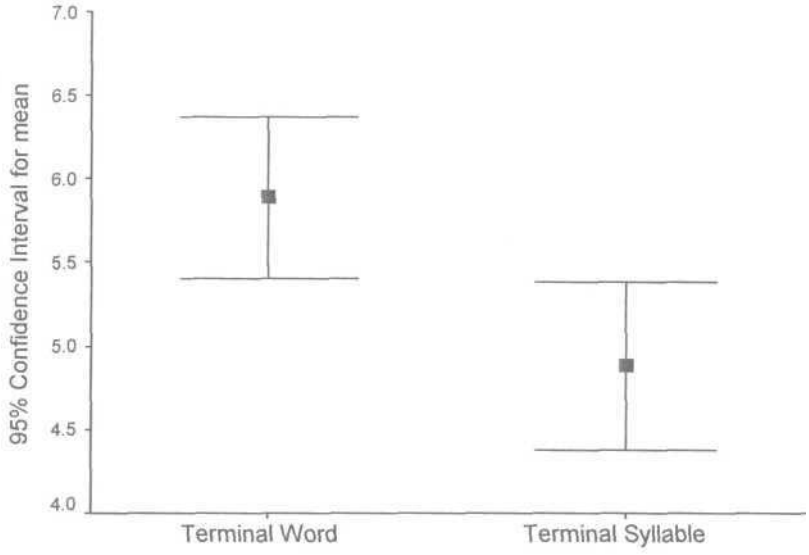
TWSTP - Terminal word steep pitch
TWSTP+* = Terminal word steep pitch+duration (250 ms/500ms/750ms)
TWSHP = Terminal word shallow pitch
TWSHP+* = Terminal word shallow pitch+duration (250 ms/ 500ms/ 750 ms)
TSSTP = Terminal syllable steep pitch
TSSTP+* = Terminal syllable steep pitch+duration (250 ms/ 500ms/ 750 ms)
TSSHP = Terminal syllable shallow pitch
TSSHP+* = Terminal syllable shallow pitch+ duration (250 ms/ 500ms/ 750ms)

Table 2 and Graph 1 depict the overall mean and standard deviation of Terminal word and Terminal syllable manipulated tokens. It is seen that the mean value of Terminal Word manipulated tokens is higher than Terminal syllable manipulated tokens. A paired sample t test was done to see if there was any significant difference between them. The results of paired sample t test indicates that the difference between Terminal word and Terminal syllable manipulated tokens is significant [$t(29) = 4.769$, ($p < 0.001$)].

Table 2: Mean and Standard deviation of responses of listeners for Terminal word and Terminal syllable manipulated tokens.

CONDITIONS	MEAN	STANDARD DEVIATION
Terminal Word	5.8875	1.3004
Terminal Syllable	4.8833	1.3479

Graph 1: Mean and Standard deviation of responses of the listeners for Terminal word and Terminal syllable manipulated tokens



From Table 2 and Graph 1 it is evident that there is significant difference between perception of terminal word and terminal syllable manipulated tokens. Higher mean value is obtained for Terminal word manipulated tokens. This finding suggests that the loci of F0 that distinguishes a statement from a Yes-No question in the standard dialect of Kannada may not be on the terminal syllable, but may be spread over the terminal word as perceived by normal listeners. This finding stands in contrast to earlier studies on several languages [Contini and Boe (1975) for French, Magno Caldognetto (1978) for Italian, Nishinuma (1979) for Japanese, Dascalu (1979) for Romanian]. It is probable that the changes over small segments of speech such as syllable may not be sufficient for the ear to perceive an utterance as a Yes-No question. Rather the envelope of acoustic changes in F0 and duration on a terminal word could have aided in the perception and identification of the utterances as a Yes-No question.

2. Perception of extent of rise {Steep (ST) vs. Shallow (SH) rise} of manipulated tokens:

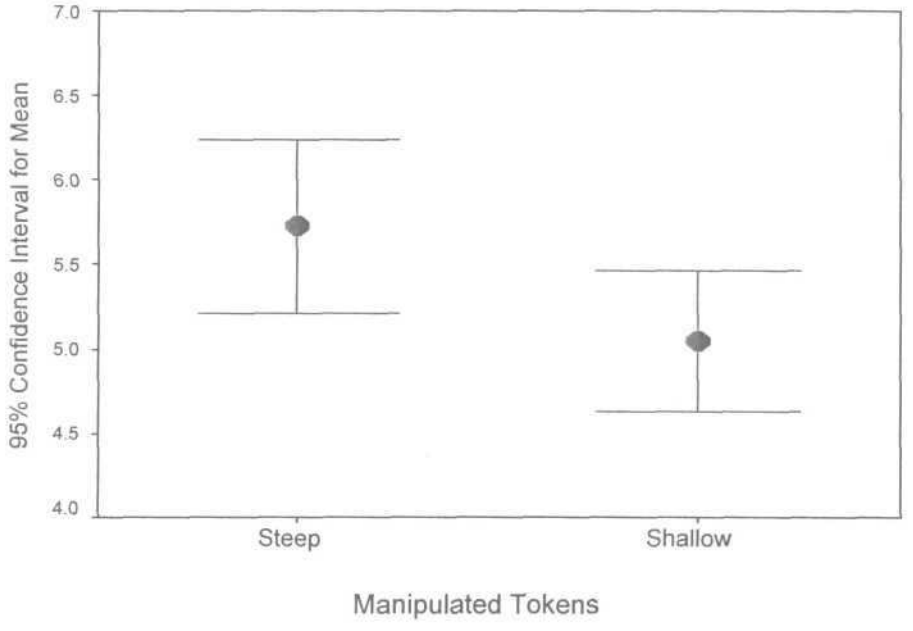
Table 3 indicates the mean and standard deviation (SD) of responses of the listener for the steep and shallow rise synthesized tokens of Terminal words and Terminal syllables. From Table 3, it can be seen that the mean values of steep rise conditions are higher than shallow rise conditions for both Terminal word and Terminal syllable manipulated tokens.

Paired sample t test was done to see if there is significant difference between the total steep (ST) rise condition of Terminal word and Terminal syllable vs. total shallow rise condition of Terminal word and Terminal syllable. Results revealed that there is significant difference between total steep rise of Terminal word and Terminal syllable vs. total shallow rise of Terminal word and Terminal syllable [$t(29) = 4.869$; $p < 0.001$]. Paired sample t test for steep and shallow conditions within terminal word manipulated tokens also indicated that there is significant difference [$t(29) = 2.668$; $p < 0.05$]. Highly significant difference between steep and shallow rise conditions for terminal syllable manipulated tokens [$t(29) = 4.252$; $p < 0.001$] was also obtained. There is also significant difference between steep rise condition of Terminal word manipulated tokens and Terminal syllable manipulated tokens [$t(29) = 3.418$; $p < 0.05$] and also between shallow rise condition of Terminal word manipulated tokens and Terminal syllable manipulated tokens [$t(29) = 4.400$; $p < 0.05$].

Table 3: Mean and Standard deviation of responses for Steep and Shallow rise condition of Terminal word and Terminal syllable manipulated tokens.

CONDITIONS		MEAN	STANDARD DEVIATION
Terminal word	Steep	6.2083	1.6957
	Shallow	5.5667	1.1725
Terminal syllable	Steep	5.2417	1.4585
	Shallow	4.5250	1.3902

Graph 2: Mean and Standard deviation of responses for Steep and Shallow rise condition of Terminal word and Terminal syllable manipulated tokens.



The results of the study show that modification of pitch height, i.e., the extent of rise (steep vs. shallow) can lead to changes in the listeners' perception of synthetic tokens. The results indicate that there is significant difference between perception of steep and shallow rise manipulated tokens. Tokens that were manipulated to a steep rise have favoured more question responses from the listeners. This shows that an increase in the extent of rise leads to increase in question judgments from the listeners. It seems that steep rise in FO cues Yes-No question intonation better than shallow rise in Bangalore-Mysore dialect of Kannada. This finding is in consonance with earlier studies which reported a major role of pitch height for utterance type identification. Ladd and Morton (1997) observed shifts in the perception of short utterance from normal to emphatic as pitch height increased. Previous studies in Hungarian language by Gosy and Terken (1994) and in Japanese by Makarova (2001) and in Russian by Rathcke (2006) also reported that a shallow rise (gradual rise) gives a smaller percentage of interrogative judgments when compared to steep rise.

3. Perception of pitch and duration characteristics of manipulated tokens

To see the effect of increase in duration superimposed on pitch altered stimuli on perception of question intonation, the pitch manipulated tokens of terminal word and terminal syllable were further increased in steps of 250 ms of duration (250ms, 500 ms, 750 ms). This was carried out in order to see if there was any difference between pitch only manipulated tokens and pitch+duration manipulated tokens. These durations were

selected by perceptually analyzing the synthetic stimuli with varying duration increase by a Speech-language pathologist well versed in the area of intonation. A minimum durational increase of 250 ms was effective in bringing about a percept of Yes-No question. Similarly manipulations above 750 ms when listened to sounded unnatural. Hence it was decided to have three levels of duration manipulation(250 ms, 500 ms and 750 ms). The raw data was analyzed and averages were calculated. Mean and standard deviation (SD) of pitch manipulated tokens and pitch+duration manipulated tokens are depicted in Table 4. It can be seen that mean values of pitch+duration manipulated tokens is higher than pitch manipulated tokens. Paired sample t test revealed significant difference between them [$t(29) = 2.048$; $p < 0.05$].

Table 4: Mean and standard deviation (SD) of responses of listeners for pitch and pitch+duration manipulated tokens:

Manipulated Tokens	**Mean**	**Standard Deviation**
Pitch (Hz)	5.1667	1.2480
Pitch+Duration (ms)	5.4583	1.2362

Graph 3: Mean and standard deviation (SD) of responses of listeners for pitch and pitch+duration manipulated tokens:

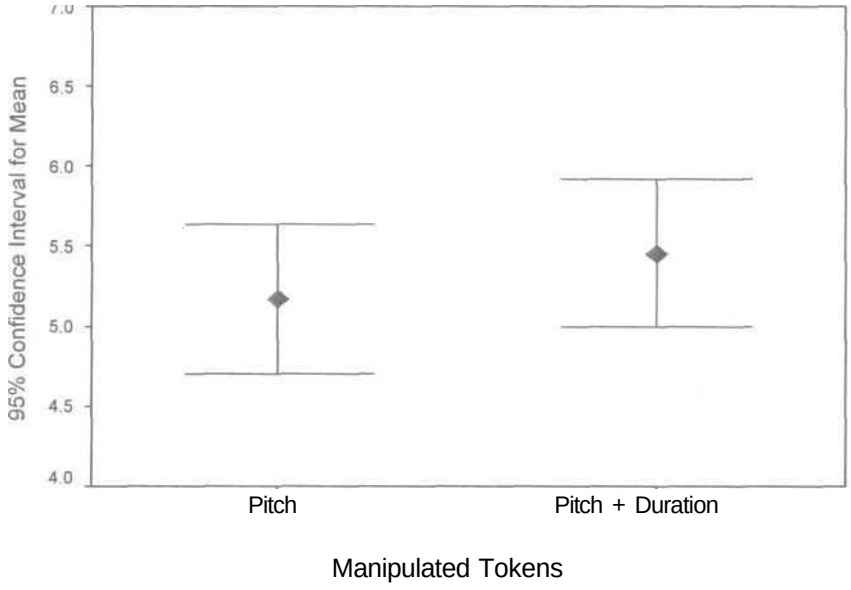


Table 5 indicates the mean and standard deviation of pitch manipulated and pitch+duration manipulated tokens in steep and shallow rise conditions of terminal word and terminal syllables. In order to see if there was any difference between the pitch manipulated and pitch+duration manipulated tokens in terminal word and terminal syllable conditions, paired t test was done. Pair wise comparisons reveal that there is no significant difference between pitch and pitch+duration manipulated tokens in steep and shallow rise conditions for both condition of terminal word and terminal syllables. The mean values however suggest that the pitch+duration manipulated tokens have favoured more question responses than pitch only manipulated tokens in all the above mentioned conditions.

Table 5: Mean and Standard deviation of responses of pitch manipulated and pitch + duration manipulated tokens in steep and shallow rise conditions of Terminal word and Terminal syllable.

CONDITIONS			MEAN	STANDARD DEVIATION
Terminal word	Steep	Pitch	6.1000	1.8448
		Pitch+duration	6.2444	1.7659
	Shallow	Pitch	5.3667	1.3515
		Pitch+duration	5.6333	1.2696
Terminal syllable	Steep	Pitch	5.0333	1.6709
		Pitch+duration	5.3111	1.4957
	Shallow	Pitch	4.1667	1.6206
		Pitch+duration	4.6444	1.5211

There was significant difference between perception of pitch manipulated tokens and pitch+duration manipulated tokens. Higher mean scores were obtained for pitch+duration manipulated tokens. This indicates that changes of pitch alone are less effective than changes of pitch+duration in the percept of Yes-No question intonation in Kannada. It shows that the listeners use a combination of prosodic components in order to make optimal perceptual judgments. This finding supports the observations made in earlier studies that duration also is a key factor to mark question statement contrast (Fry, 1955; Lieberman, 1967). Earlier studies on production of prosodic components in

Kannada indicated that both pitch and duration play a major role in bringing syllables into prominence (Manjula, 1997). From the results of the present study, it is evident that FO and duration are important cues for perception of Yes-No question intonation in Kannada.

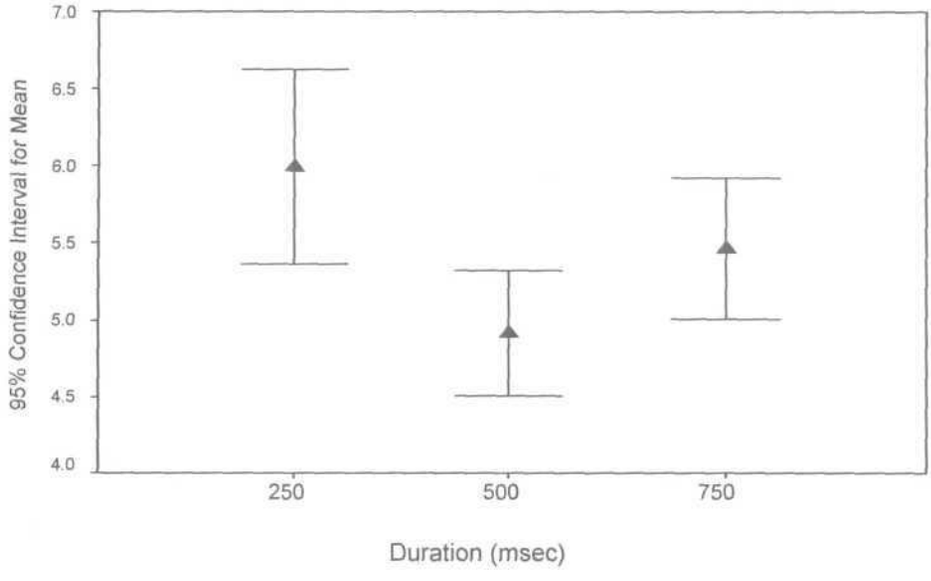
Comparison of perceptual responses among the 3 conditions of pitch+duration increase manipulation of the sentence stimuli, the duration increase was made and the same is shown in Table 6. Mean and standard deviation (SD) of perceived responses to tokens with 250 ms, 500ms and 750 ms increase in duration is represented in this table. The mean value of 250 ms increase in duration is greater followed by 750 ms increase and 500 ms increase respectively. To see if there was any significant difference between the three conditions, One way Repeated measures ANOVA was done. The results indicate that there is highly significant difference between the conditions in which duration was increased to 250 ms, 500ms and 750 ms [$F(2, 58) = 18.625$; $p < 0.001$]. Bonferroni's multiple comparison test was done to make pair wise comparisons. It also revealed that there is significant difference between the conditions. From the mean values depicted in Table 5 and Graph 4 it can be seen that the tokens in which duration was increased by 250ms was best perceived as questions followed by 750 ms increase and 500 ms increase respectively.

Table 6: Mean and Standard deviation of responses of listeners for tokens with steps of duration increase.

DURATION INCREASE	**MEAN**	**STANDARD DEVIATION**	F	P
TW+TS (250 ms)	5.9917	1.6809	18.625	0.000
TW+TS (500ms)	4.9167	1.0874		
TW+TS (750ms)	5.4667	1.2278		

Note:
1. TW+TS (250 ms) = Terminal word + Terminal syllable pitch+ duration (250 ms) manipulation.
2. TW+TS (500 ms) = Terminal word + Terminal syllable pitch+ duration (250 ms) manipulation.
3. TW+TS (750 ms) = Terminal word + Terminal syllable pitch+ duration (250 ms) manipulation.

Graph 4: Mean and Standard deviation of responses of listeners for tokens with duration increase.



Results of the present study revealed that there is significant difference in the perception of question intonation across varying increase in duration. As duration increased along with rise in pitch more number of question responses was obtained. Higher mean values were obtained for 250 ms increase in duration followed by 750 ms and 500 ms respectively. However, increase in duration beyond 250 ms has decreased the question responses by the subjects. Lower mean scores were obtained for 750 ms and 500 ms. A plausible explanation for this finding is that duration manipulation centered around 250 ms could be most optimal for perceptual identification of Yes-No questions. Any raise above this limit may not facilitate optimum responses for perception of Yes-No questions in Kannada language. The lower mean scores for 500 ms when compared to other 2 conditions could be due to individual variability in the scoring by the subjects.

4. Perception of 'naturalness' of synthesized tokens.

In order to analyse the naturalness of the synthetic tokens, the thirty listeners were asked to rate the tokens as unnatural/fairly natural or natural on the response sheet given to them. The naturalness ratings were obtained for those tokens judged as questions. The mean and standard deviation of question responses rated as unnatural are given in Table 7. One way Repeated measure ANOVA was done to see if there was significant difference between the different manipulation conditions. The results revealed there was significant difference between the conditions at $F(15,435) = 2.525$; $p < 0.05$. However, Bonferroni's multiple comparison test did not reveal any significant difference among

pair wise comparisons. This is because of the very high standard deviation (SD) values. When the mean values are compared it is evident that the mean value of Terminal word steep pitch+duration manipulation (750 ms) was the highest rated as unnatural, followed by terminal word steep pitch+duration manipulation (500 ms) and terminal word shallow pitch+duration manipulation (500 ms). Overall, Terminal word manipulated tokens have greater mean values of unnaturalness when compared to Terminal syllable manipulated tokens. This indicates that though terminal word manipulated tokens favoured higher question judgments, they were also perceived to be more unnatural when compared to terminal syllable manipulated tokens.

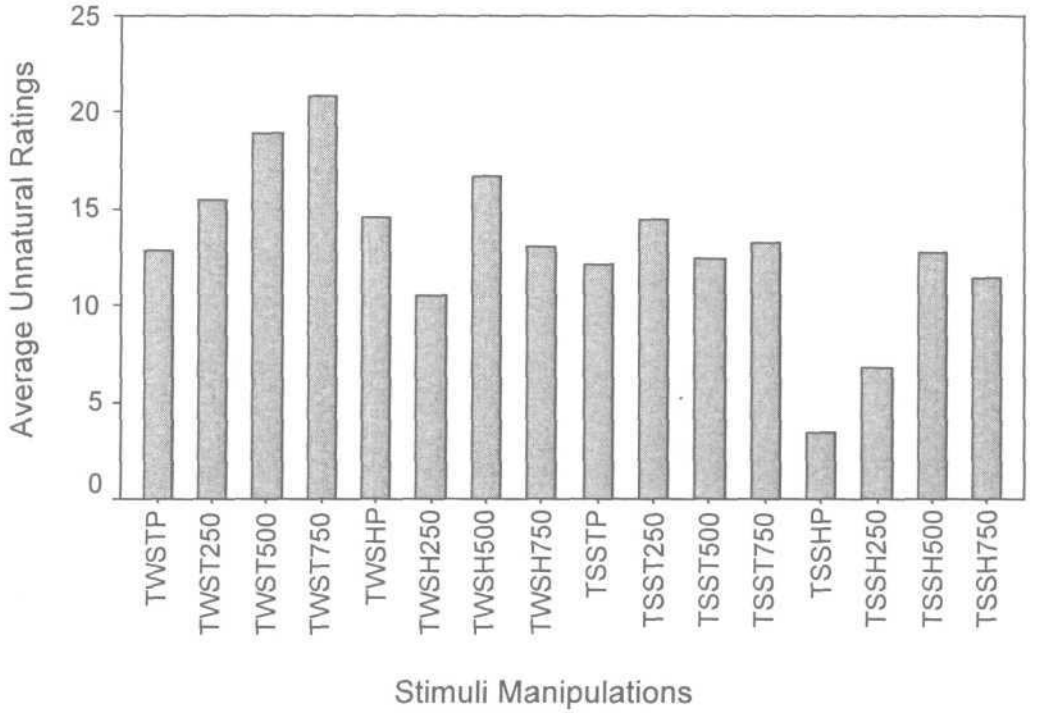
Table 7: Mean and Standard deviation of responses rated as unnatural in different manipulation conditions.

UNNATURAL TOKENS	MEAN	**STANDARD DEVIATION**
TWSTP	12.8677	21.6689
TWSTP+*250 ms	15.5083	23.5036
TWSTP+*500 ms	18.9603	25.1952
TWSTP+*750 ms	20.8677	25.8879
TWSHP ms	14.6190	22.9177
TWSHP+*250 ms	10.5040	15.3814
TWSHP+*500 ms	16.6786	25.0998
TWSHP+*750 ms	13.0079	21.6114
TSSTP ms	12.1587	22.0300
TSSTP+*250 ms	14.4365	19.3470
TSSTP+*500 ms	12.4206	23.4595
TSSTP+*750 ms	13.2937	21.4938
TSSHP ms	3.4524	8.1486
TSSHP+*250 ms	6.7593	14.4236
TSSHP+*500 ms	12.7778	26.8683
TSSHP+*750 ms	11.3889	25.7533

Note:

TWSTP = Terminal word steep pitch
TWSTP+* = Terminal word steep pitch+duration (250 ms/500ms/750ms)
TWSHP = Terminal word shallow pitch
TWSHP+* = Terminal word shallow pitch+duration (250 ms/ 500ms/ 750 ms)
TSSTP = Terminal syllable steep pitch
TSSTP+* = Terminal syllable steep pitch+duration (250 ms/ 500ms/ 750 ms)
TSSHP = Terminal syllable shallow pitch
TSSHP+* = Terminal syllable shallow pitch+ duration (250 ms/ 500ms/ 750 ms)

Graph 5: *Mean and Standard deviation of responses rated as unnatural in different manipulation conditions.*



The findings suggest that there was significant difference in the ratings of naturalness for the various manipulated tokens. Overall, Terminal word manipulated tokens have greater mean values of unnaturalness when compared to Terminal syllable manipulated tokens. This finding is interesting as it indicates that though terminal word manipulated tokens favoured higher question judgments, they were also perceived to be more unnatural when compared to terminal syllable manipulated tokens. There are no

previous studies of this nature in Kannada. The findings may be suggestive of the fact that synthetic manipulation over longer utterances tends to degrade the stimuli and thus elicit higher ratings of unnaturalness from the listeners. Synthetic manipulations over smaller segments such as a syllable would not degrade the stimuli to the extent that happen with the synthetic manipulations over terminal word.

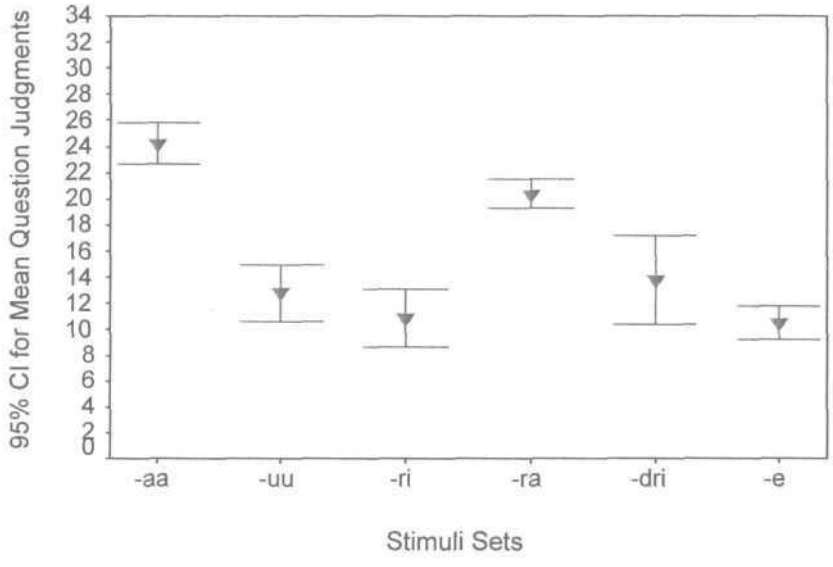
5. Perception of synthesized tokens with 6 different clitic markers:

The stimuli sentences were selected such that when the terminal manipulations of pitch and duration is done, it could be grouped into Yes- No questions with cliticization of '-aa', '-uu', '-ee', '-ra', '-ri' and '-dri'. The raw data was analysed to determine the responses as questions for 6 different clitic conditions. Table 8 indicates the mean and standard deviation (SD) for different sets of clitic conditions. One way Repeated measures ANOVA indicated that there is significant difference between the different stimuli sets [$F(5,145) = 45.190$; $p < 0.001$]. Bonferroni's multiple comparison test revealed that sentences ending with cliticization of '-aa', and '-ra' were significantly different from each other and also from other sets of stimuli. Highest mean scores were obtained for sentences ending with the cliticization of '-aa', followed by '-ra' and '-dri' respectively. This suggests that listeners perceived more number of manipulated sentences ending with the cliticization of '-aa' and '-ra' as questions when compared to sentences ending with cliticization of'-uu', '-ri' and ee'. Stimuli ending with '-ee' were the least perceived as questions.

Table 8: Mean and Standard deviation of responses of listeners for different clitic markers.

STIMULI SETS	MEAN	STANDARD DEVIATION
'-aa' clitic	24.23	4.30
'-uu' clitic	12.80	5.81
'-ri' clitic	10.86	5.92
'-ra' clitic	20.36	3.03
'-dri' clitic	13.80	8.98
'-ee' clitic	10.53	3.51

Graph 6: *Mean and Standard deviation of responses of listeners for different clitic markers.*



The secondary aim of the study was to investigate the perceptual response of listeners to stimuli with different cliticization. There are no studies of this nature in Kannada. The present study indicated that there was significant difference in the perception of question responses across different clitic sets. Results showed that sentences with cliticization of '-aa', '-ra', and '-dri' favored more question responses from the listeners. Least question responses were for stimuli ending with cliticization of '-ee'. A plausible reason for such a finding could be because in Kannada, Yes-No questions commonly end with '-aa', '-ra', and '-dri'. Hiremath (1961) had reported that vowels (i, e, ai, a, aa, u, uu) are always found in word final position in Kannada. In Kannada, subjects tend to produce more '-aa', '-ra', and '-dri' ending Yes-No questions when compared to other clitics and hence these tend to be in the periphery of their linguistic experience. Hence their idea of interrogativity is also associated with Yes-No questions having aa', '-ra', and '-dri' cliticization. Similarly in Kannada, more statements end with '-ee' clitic. This could explain the least mean scores of question responses for stimuli ending with '-ee' clitic. This result is consistent with Lindblom (1990) study where he reported that listeners cannot help imposing their knowledge of a certain language on the actual signal they hear.

The findings of the present study can be summarized as follows:

a) Terminal word manipulated tokens were perceived as questions better compared to terminal syllable manipulated tokens, indicating that the loci of F0, that cues a

statement from Yes-No questions in Bangalore-Mysore dialect of Kannada, may be spread over the terminal word than on terminal syllable, as perceived by listeners.

b) Steep rise manipulated tokens were perceived better as questions when compared to shallow rise manipulated tokens , indicating that the extent of rise is also an important feature that differentiates a statement from Yes-No questions in Bangalore-Mysore dialect of Kannada.

c) Higher scores were obtained for pitch+duration manipulated tokens when compared to pitch only manipulated tokens, indicating that listeners use a combination of pitch and duration cues to differentiate statement from Yes-No questions in Bangalore-Mysore dialect Kannada.

d) The results revealed that an increment of 250 ms is perceived as optimum durational increase by the listeners to differentiate statements from Yes-No questions in Bangalore-Mysore dialect of Kannada.

e) Another interesting finding the study has yielded is that though terminal word manipulated tokens were better perceived as questions when compared to terminal syllable manipulated tokens, it was also rated as more unnatural than terminal syllable manipulated tokens. This could probably be because of the degradation of the stimuli due to synthetic manipulations.

f) The study also revealed that sentences with cliticization of '-aa', '-ra', and '-dri' favored more question responses from the listeners when compared to other clitics. This could probably be because in Kannada, Yes-No questions commonly end with '-aa', '-ra', and '-dri'. It may be that the listeners' knowledge of interrogative structure of language has indirectly influenced the perceptual judgment of Yes-No questions having - 'aa', '-ra', and '-dri' cliticization.

SUMMARY AND CONCLUSIONS

Prosody is an important aspect of speech in human beings. Prosody may serve as the interface between low level segmental information and higher level grammatical structures in speech. It still remains the less studied and highly complex linguistic feature. Intonation, an aspect of prosody, forms the melody of speech. It is the variation of fundamental frequency (FO) or pitch with reference to a time scale.

Kannada is one of the major Dravidian languages with approximately seven dialects (Rajapurohit, 1982). The intonation and other prosodic features seem to vary from dialect to dialect. These have not been studied in Kannada. There have been few studies on production aspects of intonation in Kannada, but no study on the perceptual aspect of intonation in Kannada. Hence the present study was planned with the objectives of describing the perceptual aspects of terminal contour of question-statement intonation in Bangalore-Mysore dialect of Kannada.

The study aimed a) To determine the loci of FO change (FO over terminal syllable or terminal word) that differentiates question from statement in the Bangalore-Mysore dialect of Kannada using synthesized speech tokens subjected to perceptual judgment, b) To find the extent of rise(gradual/steep/ level) that would effectively cue statement as question, c) To find if the duration characteristics (lengthening of terminal syllable or terminal word) would effectively cue a statement from Yes- No question. As a secondary

goal the naturalness ratings of the synthetic tokens was also analysed and the perceptual response of listeners to stimuli with different cliticization was also investigated.

Subjects in the study included one model speaker and thirty normal listeners (15 males and 15 females) who were native speakers of Bangalore-Mysore dialect of Kannada. The sentence stimuli were designed such that with change in intonation they would be perceived as Yes-No questions. The model speaker was asked to utter the sentences as declarative statement. The sentence stimuli spoken by the model speaker served as the basis for preparing manipulated tokens. The terminal words and terminal syllables of these sentences were manipulated for pitch (steep rise and shallow rise) and duration (250 ms, 500 ms, 750 ms). The stimuli manipulated for pith and duration changes in the terminal words and syllables were randomized and presented to thirty listeners. The listeners were asked to judge the manipulated tokens as question or statement and also rate how natural they perceived the tokens to be on a 3 point rating scale (unnatural/fairly natural/natural). These perceived responses of the listeners to the manipulated tokens were calculated, which formed the raw data. The perceptual responses of the listeners to different manipulations were analysed. The results were discussed and inferences were drawn based on suitable statistical procedures. The salient findings are summarized as follows:

1. There was significant difference between perception of terminal word and terminal syllable manipulated tokens. Higher mean value was obtained for Terminal word manipulated tokens. This finding suggests that the loci of F0 that

distinguishes a statement from a Yes-No question in the standard dialect of Kannada may not be on the terminal syllable, but may be spread over the terminal word as perceived by normal listeners.

2. The results indicated that there was significant difference between perception of steep and shallow rise manipulated tokens. Tokens that were manipulated to a steep rise favoured more question responses from the listeners. It was observed that an increase in the extent of rise leads to increase in question judgments from the listeners. Hence, steep rise cued Yes-No question intonation better than shallow rise in Bangalore-Mysore dialect of Kannada.

3. There was significant difference between perception of pitch manipulated tokens and pitch+duration manipulated tokens. Higher mean scores were obtained for pitch+duration manipulated tokens. This finding suggests that changes of pitch alone were less salient than changes of pitch+duration in marking Yes-No question intonation in Kannada. This finding is reasoned on the basis that listeners probably use a combination of prosodic components in order to make optimal perceptual judgments. Hence in Bangalore- Mysore dialect of Kannada, both pitch and duration acts as a cue to perceive questions.

4. There was a significant difference in the perception of question intonation with manipulated tokens of duration. As duration increased along with rise in pitch more number of question responses was obtained. Higher mean values were

obtained for 250 ms increase in duration followed by 750 ms and 500 ms respectively suggesting that 250 ms increase in duration seems to be optimal to bring about perception of Yes-No questions in Kannada.

5. There was significant difference in the ratings of naturalness for the various manipulated tokens. Overall, Terminal word manipulated tokens were more often rated as "unnatural" when compared to Terminal syllable manipulated tokens. It implied that although terminal word manipulated tokens favoured higher question judgments, they were also perceived to be more unnatural when compared to terminal syllable manipulated tokens. It may be attributed to the general fact that synthetic manipulation over longer utterances tends to degrade the stimuli and thus elicit higher ratings of unnaturalness from the listeners.

6. Results showed that sentences with cliticization of '-aa', '-ra', and '-dri' favored more question responses from the listeners. Stimuli ending with cliticization of 'ee' were least rated as question responses. A plausible reason for such a finding could be because in colloquial Kannada, Yes-No questions commonly end with '-aa', '-ra', and '-dri' and statements end with '-ee". This finding suggests that listeners may impose their linguistic experience and knowledge on the actual signal they hear.

Conclusions:

In view of the above results, following conclusions can be drawn:

- The perceptual loci of FO change that differentiates a statement from Yes-No questions, is spread over the terminal word in Bangalore-Mysore dialect of Kannada.
- The extent of rise (steep vs. shallow) is also important in effectively cuing a statement as a question in Bangalore-Mysore dialect of Kannada.
- The results also suggest that the perception of Yes-No question is linked with a combination of changes in fundamental frequency (FO) and duration features in Bangalore-Mysore dialect of Kannada.
- The synthetic manipulation over larger segment of an utterance (terminal word) tends to reduce the naturalness of the percept of question intonation as compared to shorter segment of an utterance (terminal syllable). Hence the naturalness of the stimuli is also an important factor to be considered in perceptual experiments.
- The listeners' responses to cliticization of stimuli, is influenced by their linguistic knowledge of the language.

Recommendations:

- The study any be replicated in other dialects of Kannada.
- The study may be replicated by manipulating other prosodic features of interrogatives in Kannada.
- The interaction of listeners' linguistic experience on their percept can be studied in depth.

REFERENCES

Abe, I. (1955). Intonational pattern of English and Japanese. In D.L. Bolinger (Ed), *Intonation — selected readings.* England: Penguin Books Ltd.

Armstrong, L. E., & Ward, I. C. (1926). *A Handbook of English Intonation.* Cambridge: Heffer.

Austin, J. (1975). *How we do things with words.* Sbias M, Urmsson J, editors. Cambridge, MA: Harvard University Press.

Beckman, M. E. (1996). The parsing of prosody. *Language and Cognitive Processes* 11 (1): 17-67.

Brazil, D., Coulthard, M., & Johns, C. (1980). *Discourse Intonation and Language Teaching.* London: Longman.

Brazil, D. (1995). *A Grammar of Speech,* Describing English Language. Oxford: Oxford University Press.

Boersma, P., & Weenik, D. (2000). *Praat, A system for doing phonetics by computer,* version 3.4(Technical Report 132). Institute of Phonetic sciences of the University of Amsterdam.www.praat.org.

Bolinger, D.L. (1972). *Intonation - Selected readings.* England: Penguin Books Ltd.

Bolinger, D.L. (1978). Cited in *Intonation* by Cruttenden, A. (1986). Cambridge: Cambridge University Press.

Chafe, W. (1993). Prosodic and functional units of language. In J. A. Edwards and M. D. Lampert (eds), *Talking Data: Transcription and Coding Methods for Language Research.* Hillsdale, NJ: Lawrence Erlbaum

Chang (1958). In *Intonation, Perception and Language* by Lieberman, P. (1967). Cambridge: The MIT Press.

Chomsky, N. (1970). Deep structure, surface structure and semantic interpretation. In Semantics by Steinberg, D.D., and Jakobvits, L.A. (Ed.). Cambridge: Cambridge University Press.

Chun, D. M. (2002). *Discourse Intonation in L2: From theory and research to practice.* Amsterdam/ Philadelphia: John Benjamins Publishing Company.

Collier, R. (1991). On the perceptual analysis of intonation. *Speech communication,* 9, 443-451.

Contini & Boe (1975). Cited in *Prosody: Models and Measurements* by Cutler, A. and Ladd D. R. (Ed.). (1983). Berlin: Springer- Verlag.

Cooper, W., & Sorensen, J. (1981). *Fundamental Frequency in Sentence Production.* New York: Springer -Verlag.

Couper-Kuhlen, E. (1986). *An introduction to English prosody.* Edward Arnold, London.

Couper-Kuhlen, E., & Selting, M. (1996). Towards an interactional perspective on prosody and a prosodic perspective on interaction. In Couper-Kuhlen, E. and Selting, M. (eds), *Prosody in Conversation,* 11-56. Cambridge: Cambridge University Press.

Cruttenden, A. (1981a). Falls and rises: meanings and universals. *Journal of Linguistics,* 17 77- 92.

Cruttenden, A. (1986). *Intonation.* Cambridge: Cambridge University Press.

Cruttenden, A. (1992). The origins of nucleus. *Journal of the International Phonetics Association,* 20,1-9.

Cruttenden, A. (1997). *Intonation.* 2nd edition. New York: Cambridge University Press.

Crystal, D. (1969). *Prosodic Systems and Intonation in English.* Cambridge: Cambridge University Press.

Cutler, A., Dahan, D., & Donselaar, W. (1997). Prosody in the comprehension of spoken language: a literature review. *Language & Speech,* 40, 141-201.

Cutler, A., & Ladd, D. R. (1983). *Prosody: Models and Measurements.* Berlin: Springer-Verlag.

Danes, (1960). Cited in *Intonation, Perception and Language* by Lieberman, P. (1967). Cambridge: The MIT Press.

Dascalu, L. (1979). On the intonation of questions in Romanian: The rising pattern. *Review of Roum, Linguist,* 24, 35-41.

Delattre, P. (1966). A Comparison of syllable length conditioning among languages. *International Review of Applied. Linguistics,* 4,183-198.

Denes, P. (1959). Cited in Linguistic modality effects on fundamental frequency in speech by O'Shaughnessy and J. Allen. *Journal of Acoustical Society of America,* 74(4), 1155-1171, 1983.

Dubois, J.W., Cumming, S., Schuetze-Coburn, S., & Paolino, D. (1992), *Discourse Transcription,* Santa Barbara Papers in Linguistics, 4, Santa Barbara, CA, Department of Linguistics, Santa Barbara: University of California.

Dubois, J.W., Schuetze-Coburn,S., Cumming, S., & Paolino, D. (1993). Outline of discourse transcription, *in* Edwards, J. A. and Lampert, M. D. (eds.). *Talking Data: Transcription and Coding in Discourse Research,* Hillsdale, New Jersey: Lawrence Erlbaum Associates, pp 45-89.

Eady, S. J., & Cooper, W.E. (1985). Speech Intonation and Focus Location in Matched Statements and Questions. *Journal of Acoustical Society of America,* 23, 765-769

Eady, S. J., & Cooper, W.E. (1986). Speech intonation and focus location in matched statements and questions. *Journal of the Acoustic Society of America,* 80 (2), 1101-1123.

Freeman, F.J. (1983). Cited in pp 48-50 *"Intonation"* by Cruttenden, A. (1986). New York: Cambridge University Press.

Fry, D. B. (1955). Duration and intensity as physical correlates of linguistic stress. *Journal of Acoustical Society of America,* 27, 765-8.

Fry, D. B. (1958). Experiments in the Perception of Stress. *Language andSpeech,* 1, 126-152.

Garding, E., & Abramson, A.S. (1965). A study of the perception of some American English intonation contours. *Studia Linguistica,* 19, 61-79.

Gosy, M., & Terken, J. (1994). Questions marking in Hungarian Timing and height of pitch peaks. *Journal of Phonetics.* 22,269-281.

Green, M.C.L. (1964). *The voice and its disorders.* London: Pittman Medical Publishing, Co.Ltd.

Hadding-Koch, K. (1961). Acoustico-phonetic studies in the intonation of southern Swedish. Technical report, Gleerup, C. W. K., Lund, Sweden.

Hadding, K., & Naucler, K. (1980). Permissible and impermissible variations in pitch contours. In.Waugh L.R and.von Schooneveld C.H, *Melody of Language- intonation and Prosody* (Ed.). Maryland: University Park Press.

Hadding-Koch, K., & Studdert-Kennedy, M. (1964). An experimental study of some intonation contours, *Phonetica,* 11, 175-185.

Halliday, M.A.K. (1967). *Intonation and grammar in British English,* Mouton and Co (1967). The Hague: N.V. Publishers.

Hiremath, R.C. (1961). *The structure ofKannada.* Dharward: Karnatak university.

Hirst. D., & Di Cristo. A. (1998). A Survey of intonation systems In Hirst, D., and.Di Cristo, A. (eds.) *Intonation systems,* Cambridge: Cambridge University Press.

Jakobson, Roman, & Halle,.M. (1956). *Fundamentals of language.* Mouton and Co. (1956). The Hague: N.V. Publishers.

Jones, (1957). Cited in Research on intonation during the past 10 years by K.Magdics, (1963). *Acta Linguistica,* 13, 113-165.

Kingdon, (1958). Cited in Research on intonation during the past 10 years by K. Magdics, (1963). *Acta Linguistica,* 13, 113-165.

Klatt, D. (1976). Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *Journal of Acoustical Society of America,* 59, 1208-21.

Kohler, K. J., & Gartenberg, R. (1991). The perception of accents: F0 peak height versus F0 peak position. Arbeitsberichte des Instiruts fr Phonetik der Universitt Kiel (AIPUK), 25, 219-242.

Labo, (1961). Cited in Nataraja, N.P. (1981). Intonation in Four Indian Languages under five emotional conditions. *Journal of All India Institute of Speech and Hearing,* 12, 22-28.

Ladd, D. R. (1983). Phonological features of intonational peaks. *Language,* 59, 721-759.

Ladd, D.R., Silverman, K.E.A., Tolkmitt, F., Bergmann, G., & Scherer, K. R. (1985).Evidence for the independent function of intonation, contour type, voice quality, and F0 range in signaling speaker affect. *Journal of Acoustical Society of America,* 78, 435-444.

Ladd, D. R. & Morton, R. (1997). The perception of intonational emphasis: continuous or categorical? *Journal of Phonetics,* 25, 313-342.

Lee, W.R. (1980). A point about the rise endings and fall endings of yes - no questions in *The melody of language- Intonation and Prosody* by Waugh, L.R. and.von Schooneveld, C.H. (ED.). Maryland: University Park Press.

Lehiste, I. (1970). *Suprasegmentals.* Cambridge: MIT Press. .

Lehiste, I. (1976). Cited in *Prosody: Models and Measurements* by Cutler, A., and Ladd, D.R.1983. (Ed.). Berlin: Springer-Verlag..

Lieberman, P. (1959). Some Acoustic Correlates of Word Stress in American English. *Journal of the Acoustical Society of America,* 32, 451-454.

Lieberman, P. (1967). *Intonation, Perception and Language. Research Monograph,* Cambridge: The MIT Press.

Lieberman, M. Y., & Pierrehumbert, J. B. (1984). Intonational invariance under changes in pitch range and length. In Aronoff, M.,and Oehrle, R., editors, *Language Sound Structure,* 157-233. Cambridge -Massachusetts: M.I.T. Press.

Lindblom, B. (1968). Temporal organization of syllable production. Speech Transmission Laboratory (Stockholm). *Quarterly Progress and Status Report,* 2, 1-5.

Maeda, S. (1976). *A Characterization of American English Intonation.* Unpublished PhD thesis. Cambridge: MA :MIT.

Magdics, K. (1963). Research on intonation during the past 10 years. *Acta Linguistica,* 13, 133-165.

Magno Caldognetto, (1978). Cited in *Prosody: Models and Measurements* by.Cutler, A., and Ladd, D.R.(1983) (Ed.). Berlin: Spinger-Verlag.

Majewski, W., & Blasdell, R. (1969). Influence of fundamental frequency cues on the perception of some synthetic intonation contours. *Journal of the Acoustical Society of America,* 45, 450-457.

Makarova, V. (2001). Perceptual correlates of sentence-type intonation in Russian and Japanese. *Journal of Phonetics,* 29 (2), 137-154.

Manjula, R. (1979). *Intonation in Kannada- some prosodic aspects.* Unpublished dissertation, University of Mysore.

Manjula, R. (1997). *Aspects of Kannada intonation in discourse (Interrogatives).* Unpublished Doctoral thesis, University of Mysore.

Mickey, K. (1977). On 'high'and 'low' in English. *Language & Speech,* 20(1), 20-39.

Minifie, F. (1973). Cited in Nandini. (1985). Some Prosodic Aspect in Kannada. Unpublished Dissertation, University of Mysore.

Monrad-Krohn, H. G. (1947). Dysprosody or altered 'melody' of language. *Brain,* 70, 405-415.

Mouliner, E., & Charpentier, F. (1990). Pitch- synchronous waveform processing techniques for text-to-speech synthesis using diphones. *Speech Communication,* 9, 453-467.

Nandini, H.M. (1985). *Some Prosodic Aspects in Kannada.* Unpublished Dissertation. University of Mysore.

Nataraja, N.P. (1981). Intonation in four Indian languages under five emotional conditions. *Journal of All India Institute of Speech and Hearing,* 12, 22-28.

Netsell, R. (1973). Speech physiology. In. Minifie,F., Hixon, T. J., & Williams, F. (Ed.), *Normal aspects of Speech, hearing, and language.* Engelwood Cliffs, NJ: Prentice-Hall.

Schiffman, H. (1979). *A reference grammar of spoken Kannada.* Final report OE, International studies branch, U.S. Department of health, Education and Welfare.

Schmerling, S.F. (1976). *Aspects of English sentence stress.* Austin: University of Texas.

Searle, J. R. (1969). *Speech Acts. An Essay in the Philosophy of Language.* Cambridge: Cambridge University Press.

Shankarabhatta, D.N. (1978). *Kannada vakyagaLu-aantarika racane mattu artavyavasthe* Mysore: Mysore Printing and Publishing House.

Silverman, K. (1986). FO segmental cues depend on intonation: The case of the rise after voiced stops. *Phonetica,* 43, 76-91.

Sridhara, S.N. (1990). *Kannada.* London: Rouledge.

Studdert-Kennedy, M., & Hadding, K. (1973). Auditory and linguistic processes in the perception of intonation contours, *Language & Speech,* 16, 293-313.

Studdert-Kennedy, M., & Hadding, K. (1974). Are you asking me, telling me or talking to yourself? *Journal of Phonetics.* 2, 7-14.

Sweet, H. (1982). *A Power of Phonetics.* Claredon Press. Oxford, 1982.

t' Hart, J., & Cohen, A. (1973). Intonation by Rule: A Perceptual Quest, *Journal of Phonetics,* 1, 309-327.

t' Hart, J, Cohen & Collier, R. (1990). *A Perceptual Study of Intonation: an Experimental-Phonetic Approach to Speech Melody.* Cambridge: Cambridge University Press.

Tench, P. (1996). *The intonation systems of English.* London: Cassell.

Terken, J. (1991). Fundamental frequency and perceived prominence of accented syllables. . *Journal of the Acoustical Society of America,* 89(4), 1768-1776.

Thorsen, N. (1978). On the identification of selected Danish intonation contours. *Annual report of the Institute of Phonetics.* University of Copenhagen, 12, 17-73.

Thorsen, N. (1980). Cited in *Prosody: Models and Measurements* by Cutler, A. and. Ladd D.R (1983) (Ed.). Berlin: Springer-Verlag.

Uldall, E. T. (1962). Ambiguity: Question or statement? or Are you asking me or telling me?, *Proceedings of the Fourth International Congress of Phonetic Sciences* (Mouton, The Hague), pp. 779-783.

Nishinuma (1979). Cited in *Prosody: Models and Measurements* by Cutler, A.and.Ladd, D.R. 1983. (Ed.). Berlin: Springer-Verlag,

Nolan, F. (1995). The effect of emphasis on declination in English intonation. In: Windsor Lewis, J. (ed), *Studies in General and English Phonetics. Essays in Honour of Professor J. D. O'Connor,* 241-254. London: Routledge.

O'Connor, J. D. (1973). *Phonetics.* Harmondsworth: Penguin.

Ohala, JJ. (1984). An ethologial perspective on common cross-language utilization of FO of voice. *Phonetica,* 41, 1-16.

O'Shaughnessy, (1979).Linguistic features in fundamental frequency patterns. *Journal of Phonetics,!,* 119-145.

O'Shaughnessy & Allen, J. (1983). Linguistic modality effects on fundamental frequency in speech. *Journal of the Acoustical Society of America,* 74(4), 1155-1171.

Patel, R. (2002). Prosodic Control in Severe Dysarthria: Preserved Ability to mark the Question-Statement Contrast. *Journal of Speech Language and Hearing Research,* 45, 871-878.

Patil, C. (1984). Intonation contours of Kannada in *Acoustic studies in Indian languages* by B.B.Rajapurohit (Ed.) Central Institute of Indian Languages, Mysore.

Pell, M. D. (2001). Influence of emotion and focus location on prosody in matched statements and questions. *Journal of the Acoustical Society of America,* 109, 1668-1680.

Pike, K.L. (1945). *The Intonation of American English.,* USA: University of Michigan Press.

Rajapurohit, B.B. (1982). *Acoustic Characteristics of Kannada.* Central Institute of Indian Languages, Mysore.

Rathcke, T. (2006). Relevance of fO-peak shape and alignment for the perception of a functional contrast in Russian. *Proceedings of Speech Prosody 2006, I,* Dresden, pp. 65-68.

Ratna, N., Nataraaja, N.P., & Subramanyaiah, M.G. (1981). A study of prosodic aspects of Kannada language. *Journal of All India Institute of Speech and Hearing.* 12, 1-6.

Rietveld, A. C. M. & Gussenhoven, C. (1985). On the relation between pitch excursion size and pitch prominence. *Journal of Phonetics* 13, 299-308.

Ryalls, J., Dorze, G., Lever, N., Oulett, L., & Larfeuil, C. (1994). The effects of age and sex on speech intonation and duration for matched statements and questions in French. *Journal of the Acoustical Society of America,* 2274—2216.

Uldall, E. (1964). *Dimensions of meaning in Intonation* by D.Bolinger (Ed.). Penguin Books Ltd., England, 1972.

Umeda, & Quin, A.M.S. (1981).Word duration as an acoustic measure of boundary perception. *Journal of Phonetics,* 9, 19-28.

Vaissiere, J. (1983). Language Independent prosodic features in *Prosody: Models and Measurements* by A. Cutler and D. R. Ladd (Ed.), Springer- Verlag, Berlin, 1983.

von Essen (1956). Cited in *Research on Intonation during the past 10 years* by K.Magdics, (1963). *Acta Linguistica,* 13, 133-165.

APPENDIX A

SENTENCE STIMULI USED IN THE STUDY

1. Sentences with clitic "-aa"

a) ನೀನು ಈ ಬಾಗಿಲನ್ನ ತೆಗೀತೀಯ./ neenu ii baagilanna tegiitiiya
b) ಈಗ ನೀನು ಸ್ಕೂಲಿಗೆ ಹೋಗುತ್ತೀಯ./ iiga neenu skoolige hoguttiiya

2. Sentences with clitic "-uu"

a) ಈ ಚಿತ್ರ ನೀನೇ ಬರೆದಿದ್ದು. / ii citra neene barediddu

b) ಈ ಪುಸ್ತಕಾನೇ ನೀವು ತರಕ್ಕೆ ಹೇಳಿದ್ದು./ ii pustakaane neevu tarakke heLiddu

3. Sentences with clitic "-ri"

a) ನಾಳೆ ನನಗೆ ಮಿಠಾಯಿ ಕೊಡುತ್ತೀರಿ./naaLe nanage miTayi koDutiiri
b) ನಾಳೆ ನೀವು ಬರುತ್ತೀರಿ. /naaLe neevu baruttiiri

4. Sentences with clitic "-ra"

a) ಈಗ ನೀವು ತಿನ್ನುತ್ತೀರ. / iiga neevu tinnuttira
b) ನೀವು ಈ ಕೆಲಸ ಮಾಡುತ್ತೀರ. / neevu ii kelasa maaDuttira

5. Sentences with clitic "-dri"

a) ನೆನ್ನೆ ನೀವು ಹಾಡು ಹಾಡ್ದಿದ್ರಿ. / nenne neevu haaDu haaDidri
b) ನೀವು ನೆನ್ನೆ ನಮ್ಮ ಮನೆಗೆ ಬಂದಿದ್ರಿ. /neevu nenne namma manege banddidri

6. Sentences with clitic "-ee"

a) ನಾಳೆ ಹಬ್ಬ ಇದೆ. / naaLe habba ide

b) ಅವನು ಬೊಂಬಾಯಿಗೆ ಹೋಗಿದ್ದಾನೆ. / avanu bombaayige hoogiddaane

Note: Transcription based on Schiffman (1979)

APPENDIX B

RESPONSE SHEET

Name:
Age/ Sex:
Native Language:
Education:

Instruction: Now you are going to hear some recorded sentences. You have to identify whether these sentences are statement or a question and put a ( ) mark under the column for statement or question. Also, you have to indicate if the sentence heard was unnatural or fairly natural or natural by putting a ( ) mark under the respective column.

TOKEN NO	STATEMENT	QUESTION	UNNATURAL	FAIRLY NATURAL	NATURAL
1					
2					
3					
4					
5					
204					