Effect of Speaking Rate on Voice Onset Time in Telugu

Madhu Sudarshan Reddy. B. B.Sc., M.Sc. (Speech-Language Pathology)
Mahendra Kumar. N. M.Sc. (Speech-Language Pathology)
Sreedevi. N. M.Sc. (Speech & Hearing), Ph.D. (Speech-Language Pathology).

Introduction

Speech is a system of verbal communication and is distinctive to human beings. Stetson (1928) suggested that “speech is movement made audible”. In simple, audible sounds are produced by effort of vocal folds which is further altered by the articulators (tongue, lips) to generate speech. Acoustics deals with the study of sounds. The analysis of sounds in terms of temporal and spectral aspects is called acoustic analysis. It allows the speech pathologist to infer a great deal about the movement and placement of the articulators during the production of both normal and abnormal speech. Sounds comprise of vowels and consonants.
Vowels and Consonants

Vowels convey maximum energy and help in understanding speech. Consonants convey less energy but have meaningful messages in speech communication. Stops are abundantly represented in the world’s languages and often are frequently occurring consonants in a language. The stop consonants are produced by the complete occlusion of the oral cavity by articulators. Acoustic characteristics of stops include closure duration, voice onset time (VOT), release burst and formant transition. VOT is a strong cue to voicing differences between stops (Lisker & Abramson 1964, 1970). Voice onset time (VOT) is the time between the release of a stop consonant occlusion and the onset of the vocal fold vibration and is measured as the time interval between the release burst and the first quasi periodicity in the acoustic signal (Lisker & Abramson, 1964; Keating, 1984; Klatt, 1975).

Voice Onset Time

Voice onset time is influenced by language (Lisker & Abramson, 1964, 1967 in American English) and also by age (Menyuk & Klatt, 1975) and gender (Ravishankar, 1981). Most of the studies focused on two categories such as speaker related and non-speaker related factors. Extensively studied are speaker related factors such as gender (Whiteside & Irving, 1998, 2001; Ryalls, Zipper & Baldauff 1997, 2002), age (Ryalls et al., 1997), speaking rate (Kessinger & Blumstein, 1998; Volatis & Miller, 1992), lung volume (Ryalls et al., 2002) and ethnic background (Ryalls et al., 1997). The non-speaker related factors include place of articulation, word frequency, phonetic context (Whiteside et al., 2004; Neiman et al., 1983). In CV syllables, VOT is less in the context of low vowels /a/ and /æ/ when compared to high vowels /u/ and /i/ (Docherty, 1992; Higgins, Netsell, & Schulte, 1998; Klatt, 1975; Smith, 1978).

Speaking Rate

Speaking rate is typically defined as the number of output syllables or words per unit time. It depends on the speed of articular movement, degree of overlapping of co articulation and linguistic culture. Speaking rate affects the VOT values (Miller et al., 1986; Summerfield, 1975). The voiceless plosives have shorter VOTs at faster speaking rates and vice versa.
However, no significant variation is observed in the negative or short-lag VOTs of voiced plosives (Kessinger & Blumstein, 1997, 1998; Volaitis & Miller, 1992).

**Measurement of VOT**

Miller et al. (1986) measured VOT and overall syllable duration at wide range of speaking rates for /bi/ and /pi/. They found that syllable duration increased as rate was slowed and VOT values become systematically longer. Volaitis and Miller (1992) also measured the VOT for both voiceless (p, t, k) and voiced (b, d, g) stop consonants for syllable initial consonant at different speaking rates. Their results revealed that, VOT systematically increased with an increase in syllable duration. Their results replicate the findings of Miller and Baer (1986) for labial consonants and extend them to alveolar and velar consonants. Miller and Volaitis (1989) noted that, for the voiceless consonants an increase in syllable duration not only resulted in longer VOT values, but also results in a wider range of VOT values. Hence, VOT is negatively correlated with speaking rate especially for voiceless stop consonants and is highly significant (Kessinger & Blumstein, 1998; Volatis & Miller, 1992).

Robb, Gilbert and Lerman (2005) measured the VOT and syllable durations in phrases. Results revealed that females demonstrated significantly longer VOT’s for voiceless stops but not for voiced stops. Conclusion of this study was that the differences were gender based rather than speech tempo-based. Manjunath, Sneha and Narasimhan (2010) investigated the variation in VOT in Kannada language (a south Indian Dravidian Language) for both voiceless and voiced stop consonants across different speaking rates. They used the metronome with a flashlight to maintain the speaking rate and instructed the participants to read the phrases. Results revealed that slower speaking rates had higher VOT values and faster speaking rates had lower VOT values. They concluded that variation in VOT across speaking rates is attributed to the changes in dynamic specification of gesture of articulators.

Lisker and Abramson (1967) were among the first investigators to find that VOT values were influenced by the utterance type (i.e., CV syllables, CVC words, or connected speech). They concluded that voiceless plosives embedded in phrases exhibited shorter VOTs than
plosives elicited in CV syllables or CVC words. Later several studies noted the VOT across gender in isolated CV syllables and/or CVC words (Ryalls et al., 1997; Smith, 1978; Whiteside & Irving, 1997, 1998), VOT in words embedded in phrases (Robb et al., 2005; Swartz, 1992; Sweeting & Baken, 1982). Hence, variation in the utterance type across studies might relate to the variability of results.

Studies Mostly in English

Literature on VOT suggests that studies are mostly in English. Also with the discrepancies in study design and method, it is very difficult to compare the results of studies across each other. There are no studies in Telugu language reporting the effect of speaking rate on VOT across different vowel contexts and gender. Hence, the present study was planned.

Method

Participants

A total of 12 participants were considered in the present study. They were divided into two groups. Group I consisted of six adult males in the age range of 19–26 years with a mean age of 22.83 years and Group II consisted of six adult females in the age range of 19-25 years with a mean age of 21.85 years. All the participants were native speakers of Telugu language and were devoid of any structural or functional abnormalities.

Materials

A set of 18 meaningful words in Telugu language with initial stop consonants /p/, /t/, /k/, /b/, /d/, /g/ in the vowel contexts /a/, /i/, /u/ were considered (Appendix 1) and these 18 target words were embedded in a target carrier phrase: [/ippudu/ ________ /əne/ /pəḍəm/ /ʧepṱənu/] (i.e. Now I will tell the word ____ ). Thus, each participant was made to produce a total of 54 target sentences (18 sentences each in slow, normal and fast rate). A total of 648 target phrases were recorded from all the 12 participants. The recorded samples were subjected to perceptual evaluation by three judges who rated the speech as slow or normal or fast. No objective measures were used to determine the rate of speech.
Recording Procedure

Initially the participants were familiarized with the test stimuli. Following the familiarization task, the participants were instructed to utter the sentences at different rates with comfortable loudness level and pitch. All the recordings were completed in a quiet chamber and the participants were seated comfortably. The recordings were carried out using a high-quality Headset with Microphone (iBall342MV). The microphone was kept at a distance of 5-10 cm. from the mouth of the speaker. Each sentence stimuli was elicited in slow, normal and fast rate of speech consecutively. The three speaking rates were demonstrated to the participants before the recording session.

Perceptual analysis

Both inter and intra judge reliability was carried out. Three well qualified post graduates native Telugu speaking speech language pathologists who were blind to the aims of the study served as judges. Telugu is one of the South Dravidian groups of languages (Krishnamurti, 2003) and is the second most widely spoken language in India (Hussain, Durrani&Gul, 2005). Telugu language is one of the 22 official languages in India as recognized by Indian constitution in Article 343 and recently it has been declared as an official classical language (Wikimedia Foundation, 2008a). It is widely spoken in the state of Andhra Pradesh.

A total of 648 sentences recorded in slow, normal and fast rates from each participant were given to three judges for perceptual judgment of the speaking rate. The recorded tokens were randomized and each judge rated them individually as slow or normal or fast rate of speech. A high reliability co-efficient was found ($\alpha = 0.956$) after running the Cronbach’s Alpha Test. 10 % of the data were randomly selected to assess the intra-judge reliability of each judge. The results showed a high reliability co-efficiency {judge 1 ($\alpha = 0.92$); judge 2 ($\alpha = 95$) and judge 3 ($\alpha = 94$)}.

Acoustic Analysis

The recorded samples were fed into a SAMSUNG RVS509 laptop loaded with PRAAT (version 5.3.17) Software program for acoustic analysis. The recorded samples were selected randomly for analysis in PRAAT and VOT measurements were made directly from the
waveform by measuring the distance between the onset of the burst to the onset of voicing in each target syllable.

**Figure 1.** Screen display of a VOT representation analyzed in the study. The picture on the left side depicts lag VOT (positive VOT) of /pa/ and the picture on the right side depicts lead VOT (negative VOT) of /ba/ in a male speaker.

**Statistical Analysis**

Mann-Whitney U test was carried out to assess the gender differences. Since data showed significance gender difference, further statistical analysis was carried out separately for males and females. Friedman test was used to obtain the effect of speech rate on VOT. Wherever there were significant differences across the three rates, Wilcoxon’s signed rank test was used for pairwise analysis in different rates that is slow, medium and fast. Cronbach’s alpha test was used to assess the inter-judge reliability.

**Results**

The present study aimed to analyze the effect of speaking rate on VOT across different vowel contexts and gender in Telugu. There were 6 plosives including voiceless and voiced bilabials, dentals and velars. The target tokens were analyzed individually and the mean and Standard Deviation of VOT in different vowel contexts /a/, /i/ and /u/ are presented in Tables 1, 2 and 3 respectively.
Table 1: Shows the mean and standard deviation of VOT across speaking rate in the context of vowel /a/

<table>
<thead>
<tr>
<th>/a/ context</th>
<th>Slow rate</th>
<th>Normal rate</th>
<th>Fast rate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Males</td>
<td>Females</td>
<td>Males</td>
</tr>
<tr>
<td></td>
<td>(10.17)</td>
<td>(2.40)</td>
<td>(2.98)</td>
</tr>
<tr>
<td>/t/</td>
<td>22.40</td>
<td>15.56</td>
<td>22.26</td>
</tr>
<tr>
<td></td>
<td>(9.50)</td>
<td>(4.17)</td>
<td>(3.75)</td>
</tr>
<tr>
<td>/k/</td>
<td>35.28</td>
<td>32.98</td>
<td>34.98</td>
</tr>
<tr>
<td></td>
<td>(4.83)</td>
<td>(7.03)</td>
<td>(6.92)</td>
</tr>
<tr>
<td>/b/</td>
<td>113.91</td>
<td>119.50</td>
<td>72.83</td>
</tr>
<tr>
<td></td>
<td>(26.24)</td>
<td>(29.41)</td>
<td>(32.27)</td>
</tr>
<tr>
<td>/d/</td>
<td>120.08</td>
<td>86.46</td>
<td>83.96</td>
</tr>
<tr>
<td></td>
<td>(36.75)</td>
<td>(17.35)</td>
<td>(21.20)</td>
</tr>
<tr>
<td>/g/</td>
<td>101.28</td>
<td>98.71</td>
<td>96.20</td>
</tr>
<tr>
<td></td>
<td>(33.43)</td>
<td>(28.64)</td>
<td>(16.06)</td>
</tr>
</tbody>
</table>

Table 2: Shows the mean and standard deviation of VOT across speaking rate in the context of vowel /i/
Table 3: Shows the mean and standard deviation of VOT across speaking rate in the context of vowel /u/

<table>
<thead>
<tr>
<th>/u/ context</th>
<th>Slow rate</th>
<th>Normal rate</th>
<th>Fast rate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Male</td>
<td>Female</td>
<td>Male</td>
</tr>
<tr>
<td>/p/</td>
<td>31.76</td>
<td>37.25</td>
<td>31.58</td>
</tr>
<tr>
<td></td>
<td>(11.81)</td>
<td>(17.12)</td>
<td>(5.7)</td>
</tr>
<tr>
<td>/t/</td>
<td>29.83</td>
<td>20.62</td>
<td>21.06</td>
</tr>
<tr>
<td></td>
<td>(5.31)</td>
<td>(7.07)</td>
<td>(5.56)</td>
</tr>
<tr>
<td>/k/</td>
<td>40.73</td>
<td>34.90</td>
<td>38.62</td>
</tr>
<tr>
<td></td>
<td>(8.88)</td>
<td>(10.64)</td>
<td>(7.18)</td>
</tr>
<tr>
<td>/b/</td>
<td>122.55</td>
<td>119.71</td>
<td>91.73</td>
</tr>
<tr>
<td></td>
<td>(27.37)</td>
<td>(19.72)</td>
<td>(28.97)</td>
</tr>
<tr>
<td>/d/</td>
<td>132.71</td>
<td>109.65</td>
<td>91.81</td>
</tr>
<tr>
<td></td>
<td>(32.89)</td>
<td>(28.06)</td>
<td>(24.98)</td>
</tr>
<tr>
<td>/g/</td>
<td>112.11</td>
<td>109.5</td>
<td>92.28</td>
</tr>
<tr>
<td></td>
<td>(43.53)</td>
<td>(23.35)</td>
<td>(14.83)</td>
</tr>
</tbody>
</table>

It can be seen from Tables 1, 2 and 3 that the mean VOT values for voiced stops are longer than that of their voiceless cognates in all the three vowel contexts viz. /a/ /i/ /u/ and across the three rates of speech – slow, normal and fast. Thus, mean lead VOT values are longer than the lag VOT in phrases with slow, normal and fast rate of speech. Also, the mean VOT are longer for males than females in most of the tokens. It can also be observed that the VOT values
reduce as the rate of speech increases from slow to fast rate of speech in nearly all of the tokens analyzed.

**Gender Effect**

In slow rate of speech VOT values differed significantly between males and females in the context of /i/ [\(|Z| = 2.88, p < 0.05\)] and /u/ [\(|Z| = 2.08, p < 0.05\)] for unvoiced dental plosive /t/. In normal rate of speech, VOT values were significantly different between males and females in the context of /a/ and /i/ for unvoiced bilabial plosive /p/ [\(|Z| = 2.72, p < 0.05\)]; /pi/ [\(|Z| = 2.08, p < 0.05\)] and for dental plosive /t/ [\(|Z| = 2.08, p < 0.05\)]; /ti/ [\(|Z| = 2.88, p < 0.05\)]. In fast rate of speech, VOT values varied significantly between males and females for velar /k/ in the context of /a/ [\(|Z| = 2.40, p < 0.05\)], dental /t/ in the context of /i/ [\(|Z| = 2.40, p < 0.05\)] and /b/ in the context of /u/ [\(|Z| = 2.08, p < 0.05\)]. Rest of the tokens did not show any significant difference across gender.

Thus, the tokens which had significant gender differences include: /ti/ and /tu/ in slow rate; /pa/ /pi/ /ta/ and /tu/ in normal rate and /ti/, /ka/ and /bu/ in fast rate of speech. The gender differences observed were present predominantly in anterior unvoiced stops (bilabials and dentals). The mean VOT values of males were higher in all the tokens except for voiced bilabial stop /b/ in the context of /u/ in fast rate of speech.

**VOT Across Rate**

A total of six stops were considered in three vowel contexts leading to 18 target sentences. The mean VOT values reduced systematically as the rate of speech increased for all the bilabial, dental and velar stops for both voiced and unvoiced cognates and in all the three vowel contexts (Table 1, 2 & 3). The CV combinations which had statistically significant difference between the three rates of speech on Freedman test were further analyzed using Wilcoxon’s signed rank test and the results are displayed in Tables 4 and 5 for males and females respectively.

Table 4: Shows significant differences across the three rates of speech for the target CV combinations in males
## Table 5: Shows significant differences across the three rates of speech for the target CV combinations in females

<table>
<thead>
<tr>
<th>Combination of stop and vowel</th>
<th>Rate of speech</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Slow-Normal</td>
</tr>
<tr>
<td>*/tʰi/</td>
<td>*</td>
</tr>
<tr>
<td>*/bi/</td>
<td>*</td>
</tr>
<tr>
<td>*/bu/</td>
<td>*</td>
</tr>
<tr>
<td>*/dʰa/</td>
<td>*</td>
</tr>
<tr>
<td>*/di/</td>
<td>*</td>
</tr>
<tr>
<td>*/dʰu/</td>
<td>*</td>
</tr>
<tr>
<td>*/ga/</td>
<td>*</td>
</tr>
<tr>
<td>*/gi/</td>
<td>*</td>
</tr>
<tr>
<td>*/gu/</td>
<td>*</td>
</tr>
</tbody>
</table>

* indicates tokens which have significant difference at p<0.05
It can be observed from Table 4 and 5 that, out of 18 target tokens, 11 in males and 9 in females have significant differences across different rates of speech. More number of voiced tokens showed significant difference for VOT across rate compared to voiceless tokens.

Another observation on comparison of tokens across rate is that there was a significant difference between slow and fast rates and also many of the tokens vary across normal and fast rates in both males and females. However, only 2 out of 11 tokens among males and 5 out of 9 tokens among females vary significantly between slow and normal rate of speech. This indicates that the VOT values reduce with increase in rate of speech from slow to fast while the differences may not be statistically significant between slow and normal rate.

Discussion

Present study had several interesting findings. Males had higher mean VOT values than females. And also the gender differences observed were present predominantly in anterior unvoiced stops (bilabials and dentals). This can be reasoned out from an aerodynamic perspective. Some studies discussed the gender differences in VOT based on the smaller vocal tract volume among females. This size difference would result in longer VOTs in females because of the additional time required to achieve the trans-glottal pressure difference for voicing to begin (Cho & Ladefoged, 1999; Docherty, 1992). An effect of this size difference would be related to the greater accuracy of articulatory movements among females (Simpson, 2001, 2002).

It was also found that mean VOT values for voiced stops were higher than their voiceless cognates in all the three vowel contexts viz. /a/ /i/ /u/ and in all the three rates of speech – slow, normal and fast. These results are consistent with the findings in other Indian languages that the VOT for voiced stops (lead VOT) is longer than the VOT of voiceless stops (lag VOT) (Savithri & Sreedevi 1991).

The results also showed that VOT values reduce with increase in rate of speech from slow to fast rate while the differences were not statistically significant from slow to normal rate. Thus, normal to fast rate had reduced VOT when compared to normal to slow rate. Similar findings were observed in English where it was found that VOT is negatively correlated with

**Language in India** [www.languageinindia.com](http://www.languageinindia.com) ISSN 1930-2940 14:3 April 2014
Madhu Sudarshan Reddy. B. M.Sc. SLP, Mahendra Kumar. N. M.Sc. SLP, and Sreedevi. N. Ph.D.
Effect of Speaking Rate on Voice Onset Time in Telugu
speaking rate especially for voiceless stop consonants and is highly significant (Kessinger & Blumstein, 1998; Volatis & Miller, 1992).

The reduced VOT at fast rate of speech in the present study might be because of the reduced time difference between articulatory release and laryngeal release is in all voiceless cognates. All temporal parameters including VOT are also compromised. Earlier study by Manjunath et. al (2010) also reported that VOT values were higher at slower speaking rates and lower at faster speaking rates. They attributed these changes in VOT primarily to the changes in dynamic specifications of gesture of articulators. To conclude, it can be said that VOT varies with rate of speech and also with gender in Telugu. This finding augments our understanding on the physiology of speech production and also variation in sub phonemic features of plosives such as VOT.

References


=====================================================================

**Appendix 1**

List of stimuli words used in the study

/pəlli/ /pIIli/ /pullI/

/təttə/ /ṭIttu/ /ṭummu/

/kəkku/ /kIkku/ /kukkə/

/bədi/ /blddə/ /buddI/

/ðəggu/ /ḍttə/ /duduu/

/ɡəddI/ /ɡudI/ /ɡIrI/

Note: The carrier phrase used is - (/Ippudu/ _______ /əne/ /pədəm/ /ʃeptənu/)

=====================================================================

**Acknowledgement**
Authors would like to thank Prof. S. R. Savithri, Director, AIISH, for permitting us to conduct the study. Due thanks are also rendered to all participants in the study.

Madhu Sudarshan Reddy. B. M.Sc. (Speech Language Pathology)  
All India Institute of Speech & Hearing  
Mysore 570 006  
Karnataka  
India  
Madhusudarshan@rocketmail.com

Mahendra Kumar. N. M.Sc (Speech Language Pathology)  
All India Institute of Speech & Hearing  
Mysore 570 006  
Karnata  
India  
mahe15kn@gmail.com

Dr. Sreedevi. N  
Reader & Head  
Department of Speech Language Sciences  
All India Institute of Speech & Hearing  
Mysore 570 006  
Karnataka  
India  
srij_01@yahoo.co.in