International Journal of Engineering Science Invention
ISSN (Online): 2319 – 6734, ISSN (Print): 2319 – 6726
www.ijesi.org
Volume 1 Issue 1 ǁ December. 2012 ǁ PP.05-12

The Perception of Consonant Voicing and Place of Articulation in Fricatives-Vowels with Consonant Duration Modifications

N H Shobha

Professor and Head, Dept. of Electronics and Communication Engineering.
Methodist College of Engineering and Technology, Hyderabad, India

ABSTRACT: *The previous research suggests that a speaking style known as clear speech is more intelligible than casual conversational speech for hearing impaired listeners. The present paper is an extension of author's previous work addressing the temporal modifications to the clear speech intelligibility advantage. The primary goal of this work is to determine whether clear speech enhances fricative intelligibility for normal-hearing listeners with simulated impairment. Two acoustic correlates of fricative consonants namely, noise duration and formant transition duration are time-expanded selectively by 50/ 100% of their original duration. The effects of consonant duration modifications on the perception of production-based features such as voicing and place of articulation were studied. The perception in noise tests are quantified in terms of information transmission analysis measures, in the presence of white noise-masker at three noise levels, 0 dB, +12 dB, and +6 dB. Results suggested that formant transition lengthening by 50% of their original duration reported significant intelligibility improvement with respect to place of articulation feature in all three vowel contexts, in consistent with previous findings. Noise duration lengthening by 50% and 100% have less perceptual significance and does not seem to carry cues to voicing and place of articulation.*

Keywords: Clear speech, Consonant duration modification, Consonant voicing, place of articulation, Information transmission analysis, Speech intelligibility

I. INTRODUCTION

Vowel sounds carry the power in speech, but the consonant sounds are the most important for understanding. However, consonants within the same class, are often difficult to differentiate and vulnerable to signal degradation. Thus the most interesting examples of the perception of sounds is that of the perception of the consonants. Consonant perception is an area of continuing controversy for normal and impaired listeners unlike vowel perception. While the acoustic cues leading to discrimination of manner of articulation are understood, the search continues for invariant cues for consonant voicing and place of articulation.

Speakers naturally revise their speech when talking to impaired listeners or in adverse environments. This type of speech, known as clear speech, is half the typically speaking rate of conversational speech. Other differences include, more salient consonant contrasts or increased consonant-vowel ratio (CVR), consonant duration (CD), and some of the systematic changes in encoding phonetic contrasts like slower speaking rates, intense bursts, longer formant transitions, less vowel reduction[1,2,3,4]. In principle, it should be possible to improve intelligibility by exaggerating those phonetic cues not easily perceived by impaired listeners. The present study seeks to evaluate the intelligibility of synthetic clear speech, by the method of acoustic parameter alterations. The previous works of the author on CVR modification [5, 6] and CD modifications [7, 8], have evaluated CVR/CD effects on speech perception in noise. The CD modifications on plosive-vowels and fricative-vowels [7,8] , have established, (i) the perceptual comparison between voiceless and voiced consonants, and (ii) the effect of vowel contexts, in noise free and noisy environments; while the present investigation which is an extension of the above work [8] allows documenting fundamental aspects of how successfully listeners can enhance perception in terms of (i) consonant-voicing, and (ii) consonant place of articulation, of intended target fricative-vowels both in noise free and noisy environments.

Consonant and vowel phonemes are classified in terms of 'manner of articulation', 'place of articulation', and 'voicing'. Manner of articulation concerns how the vocal tract restricts airflow, completely stopping airflow by an occlusion creating stop consonant; vocal tract restrictions of varying degree creating fricatives, liquids, glides, and vowels; lowering the velum causing the nasal sounds. Place articulation refers to the location in the most narrow constriction in vocal tract usually in terms of the upper wall. It is the main attribute that enables finer discrimination of phonemes and most often associated with consonants, rather than vowels. Eight regions or points are traditionally associated with consonant constrictions as, labial, dental, alvelor, palatal, velars, uvular, pharyngeal, glottal. The phoneme is considered voiced if the vocal folds vibrate otherwise voiceless. Virtually all languages employ vowels, nasals, stops, and fricatives, but the number and choice of place of articulation within each class are highly variable across languages [9]. Fricative consonants are characterized by a turbulent noise, and may consist of the noise alone or may consist of the noise together with vocal cord vibration [10].

It is particularly unfortunate that clearly produced fricative consonants have not been the subject of more observations, since the previous consonant confusion analyses have reported that fricatives contribute a large source of errors for hearing-impaired listeners and for normal-hearing listeners in noise [11, 12]. Several studies have attempted to delineate stable acoustic correlates of the fricative voicing and place of articulation. Parameters that seem to influence identification include gross spectral shapes and peak frequencies, formant transition information, fricative noise amplitude and duration [13, 14]. The main acoustic cues that have been reported to affect perception of fricatives (for normal hearing listeners) include – Noise Duration (ND) and amplitude, as well as adjacent Formant Transition Duration (FTD) [10]. In this direction, it was decided to explore the effects of the selective time expansion of Noise Duration, and Formant Transition Duration on perception of consonant voicing and place of articulation of fricatives in the presence of noise. The work presented in this paper has two integral phases; first phase is directed towards the development of the acoustic data-base facility, and the second phase towards the perceptual tests and speech intelligibility evaluations.

II. METHOD

2.1 Participants

Three male and two females in the age group of 16-45 years with normal hearing, participated in the listening experiments. None of the subjects were experienced with perceptual experiments; subjects went through a speech token familiarization training session before the experiment started.

2.2 Speech Material

The six fricative consonants - {/f θ s v ð z/}, classified as voiced fricatives: {/v, ð, z/},voiceless fricatives: {/f θ s/)}; spanning three places of articulation : labio dental {/f, v/}, dental {/ð, θ/}, alveolar: {/s, z/}, were chosen along with three primary cardinal vowels {/a, i, u/} forming three subsets of nonsense (CV) syllables : context-/a/: {/fa,θa,sa,va,ða,za/}, context-/i/: {/fi,θi,si,vi,ði,zi/}, and context-/u/: {/fu,θu,su,vu,ðu,zu/}.

2.3 Speech Signal Processing

The speech processing protocol for development of acoustic database resembled the one employed by the author in the previous work [8]. The natural stimuli were recorded in a quiet room, sampled at 44.1K Hz, using a Praat monosound recorder. The best utterance out of 20 utterances of the author (middle aged, female) was selected based on the phonetic clarity. The speech tokens were subjected to resynthesis using the procedure of LPC (linear prediction) analysis-synthesis as provided in PRAAT [15]. The idea behind the resynthesis was two-fold; firstly, the synthetic copy renders efficient and independent manipulation of the spectral, temporal and intensity characteristics; secondly, synthetic speech is as similar as possible to a human utterance. The resynthesis tradition assumes five formants in the range between 0 to 5500 Hz for a female voice, 0 to 5000 Hz for a male voice and 0 to 10000 Hz for child voice. For implementing linear prediction with Praat, stimuli has to be band-limited by resampling the original signal to 11 KHz for female, 10 KHz for male, or 20 KHz for a young child. In the current investigation, after performing resampling at 11 KHz, the filter and the source were extracted from resampled stimuli using linear-prediction analysis. The analysis procedure adopted 10 linear-prediction coefficients (yields at most 5 formant-bandwidth pairs) in each time frame of 5 or 10 ms, which is suited for capturing changes in the speech signal. Next, using the extracted source and filter, the speech sound was regenerated based on LPC synthesis. This procedure gave back the resynthesized version with the original quality except that the windowing caused few ms at the beginning and the end of the signal to be set to zero. Finally, these tokens were normalized to 70 dB IL (referred as baseline syllables) to avoid the signal clipping in subsequent processing stages.

The consonant segment durations such as fricative noise duration and formant transition duration measurements were measured by visual inspection of the time waveforms and wideband spectrograms using the PRAAT software. 'Noise Duration- ND' is referred to the high frequency noise, measured as the difference between the fricative offset time and fricative onset time. The fricative onset time is the point at which high frequency energy appeared on spectrogram and/or point at which the number of zero crossings rapidly increased, while the fricative offset time is the intensity minimum immediately preceding the onset of vowel periodicity, for voiceless fricatives and the earliest pitch period exhibiting a change in waveform from that seen throughout the initial frication, zero crossing of the preceding pitch period was designated as fricative offset, for voiced fricatives [16]. The 'Formant Transition Durations-FTD', is the time interval between the onset of the following vowel and the instance when a formant frequency reaches its steady-state value. FTD are measured by simultaneous consultation of time domain waveform, spectrogram, linear-predictive coding (LPC) spectra, and short-time fast fourier transform (ST-FFT) spectra [17]. The LPC spectrum was constituted for a prediction order of 10 (at least twice as the number of spectral peaks that we want to detect), analysis window of 12.5 ms and 5 ms step, +6dB/octave filtering above 50 Hz. The three formants were originally located by examining the LPC spectra, FFT spectra, and spectrogram. The steady-state point of the vowel was centered at 100 ms after the onset. Formant analysis was performed for the detection of formant transition duration. After proper settings, formant contour was extracted and the formant values were written to a text file. Utilizing this data, the duration of the transitions and their onset and offset points were determined, and then applied a time warp to all formants over the determined duration of the transition. The acoustic segmentations and measurements were done using PRAAT software.

The extracted acoustic segments were subjected to time expansion using 'Pitch-Synchronous Overlap and Add' (PSOLA) algorithm. The PSOLA analysis-modification-synthesis method belongs to the general class of STFT (short-time Fourier Transform) analysis-synthesis method. The analysis phase performs the segmentation of the input speech, and the synthesis phase generates a time stretched version by overlapping and adding time segments extracted by the analysis phase. In the PRAAT object window, PSOLA can be found as sound>Convert>Lengthen (PSOLA). Here, the term 'factor' decides the factor for lengthening or shortening; by choosing factor value >1 or <1, the resulting sound could be longer or shorter than the original segment, but a factor value larger than 3 will not work. We selected a minimum pitch of 75Hz and a maximum pitch of 600Hz, while a 'factor' of 1.5 for 50% lengthening(compared to original duration) and a 'factor' of 2 for 100%lengthening(compared to original duration). Further, the lengthened segment was blended back to its original location to result in time stretched version. After the time expansion phase, duration modifications took two different schemes: (i) Noise duration modification - NDM, and (ii) Formant Transition Duration modification- FTDM. Thus, each scheme presented the stimuli at three levels of modifications: no-modification: 0%, and expanded levels of 50% and 100% (compared to original stimuli).

To simulate the hearing impairment, the acoustic dynamic range was reduced by mixing the modified stimuli with masking noise. The masking noise responsible for the threshold elevation is believed to be predominantly of cochlear origin [19]. As reported in literature, the reduction in the hearing threshold can be approximately simulated by addition of white noise [20, 21]. Some researchers have employed multi-talker babble instead of white noise [20, 21, and 22]. However, due to its non-stationary nature, the effective masking it may provide during stimulus presentation is unpredictable. Hence, we decided to use white noise masker to model the hearing loss to a good approximation. The processed tokens from the previous stage were additively mixed with the synthesized noise at three noise levels, noise-free, +12 dB SNR, and +6dB SNR. The SNR refers to the ratio of the average power in CV token to the average power of the noise token in decibels. For deriving +12dB and +6dB SNR speech tokens, the average power level of the speech token was fixed while that of the noise was adjusted. PRAAT scripts were run for synthesizing the white noise and for the process of mixing [23]. The Chris-Darwin algorithm was adopted, which performed additive mixing summed up the sounds by point-to-point values, preserving real time across the time domains [23]. Finally, after all four stages of processing stimuli corpus holds 324(162+162) test tokens spanning across 18 syllables(6*3) with 2 duration lengthening schemes(NDM,FTDM), 3 expansions(0%,50% and 100%) per scheme and 3 SNRs(noise-free, +12 dB, 6 dB) per expansion.

2.4 Speech Perception in Noise (SPIN) Tests

The SPIN tests were conducted using computerized testing procedure developed using MATLAB, where speech tokens being presented at the most comfortable listening level of 75 to 85 dB SPL for the listeners. The test procedure used a similar protocol for all three experiments. The experimentation worked on 162 tokens for each modification scheme. The subjects were played every token in ten randomized replications; were prompted to choose from the set of six choices (per vowel context) displayed on the computer screen. The perception tests were automated using a MATLAB code with graphic user interface. The results were cast as 6 *6 stimulus-response confusion matrix (CM) for a listener in a vowel context.

2.5 Speech Intelligibility Measures

The intelligibility scores of SPIN tests were summarized as the mean correct responses across listeners. The sum of the diagonal elements of CM gives the empirical probability of correct responses, known as recognition score - RS (or articulation score). The computation of RS is simple, but it obscures the detailed and important information on the distribution of errors among the off-diagonal cells [24]; also it is sensitive to the subject's bias or chance scoring (an artificially high score). Hence relative information transmission analysis approach [11], which provides a measure of covariance between stimuli and responses was adopted.

The information measures of the input stimulus X and output response Y are IS(X) and Ir(Y) respectively, are represented by the following functions:

$$I_S(X) = -\sum_i p(x_i)\log_2(p(x_i))\, bits, and \quad \text{…. (1)}$$

$$I_r(Y) = -\sum_j p(y_j)\log_2(p(y_j))\, bits \quad \text{…. (2)}$$

An MLP measure of the covariance of stimulus-response is

I(X; Y) = MLP(X) +MLP(Y)-MLP (XY) …. (3)

$$I(X;Y) = -\sum_i\sum_j p(x_i, y_j)\ \log_2\left(\frac{p(x_i)\,p(y_j)}{p(x_i, y_j)}\right) bits \quad \text{…. (4)}$$

The relative transmission from X to Y is given by,

$$Rtr(X;Y) = \frac{I\ (X;Y)}{I_s(X)} \quad \text{…. (5)}$$

Since IS(x) $\geq$ I(X;Y) $\geq$ 0; 1 $\geq$ Rtr(X;Y) $\geq$0

The above measure of relative information transmission takes into account the patterning of errors.

III. PERCEPTION OF CONSONANT VOICING AND PLACE OF ARTICULATION

To measure the transmission of information about voicing and about placing, 2*2 and 3*3 CMs were derived from the obtained 6*6 CMs and the covariance measure of intelligibility were applied to the derived sub matrices [11, 25]. This breakdown of the CM into smaller matrices and the measurement of transmission for each of these CM separately is equivalent to testing two different communication channels simultaneously [11]. The obtained transmission scores were averaged across subjects in one vowel context. The consistency of scoring pattern was being tested for statistical significance by ‘two tailed t-tests’. The statistical tables presented below reports the mean percent-correct scores, standard deviations (SD), probability value (p) and their statistical significance value. The processing factor examined the intelligibility benefit between the unprocessed speech and the processed speech, a benefit was treated significant at 0.05 levels; p<=0.01 was accepted as indicative of high significance and 0.01<p<0.05 as moderate significance.

3.1 NDM Scheme-Perceptual and Statistical Analyses

Noise duration modifications and their effects on intelligibility in terms of perceptual scores and statistical pattern are presented in Table I and II below respectively. The ND modifications resulted in varied perceptional scores with vowel context and voicing/placing feature.

Voicing: From Table I, (i) Noise-free presentations: With 50% NDM, Voicing feature benefits were +3%, +4% and 0% , while with 100% NDM, benefits were +8%, +4%, and -4% in /a/,/i/ and /u/ contexts respectively.

(ii) +12 dB SNR presentations: With 50% NDM, Voicing feature benefits were -6%, +4% and +13% ; while with100% NDM, benefits were +3%, +9% and +13% in /a/,/i/, and /u/ contexts.

(iii) +6dB SNR presentations: With 50% NDM, Voicing feature benefits were +1%,+3%, and -4%; while with 100% NDM, benefits were +4%,-7% and 0% in /a/,/i/,/u/ contexts respectively.

Placing: From Table I, (i) Noise-free presentations: with 50% NDM, Placing feature benefits were -2%, -1% and 0% ; while with 100% NDM, benefits were -2%, -1%, and -1% in /a/,/i/ and /u/ contexts.

(ii) +12 dB SNR presentations: With 50% NDM, Placing feature benefits were 0%, +1% and +9%; while with 100% NDM, benefits were +0%, +3% and +7% in /a/,/i/, and /u/ contexts.

(iii) +6dB SNR presentations: With 50% NDM, Placing feature benefits were +3%,-5%, and -1%; while with 100% NDM, benefits were +3%,+2% and -1% in /a/,/i/,/u/ contexts respectively.

From Table II, NDM scheme and its effects on statistical significance has reported no significant benefit with respect to voicing and placing feature both in the absence and presence of noise.

3.2 FTDM Scheme-Perceptual and Statistical Analyses

Formant transition duration modifications and their effects on intelligibility enhancements, both in terms of perceptual and statistical scores are presented in Table III and IV below. The FTD modifications resulted in varied perceptional scores with vowel context and voicing/placing feature in three SNR presentations.

Voicing: From Table III, (i) Noise-free presentations: With 50% FTDM, Voicing feature benefits were +5%, +3% and 0% benefits ; while with 100% FTDM, benefits were +8%, +1%, and -7% in /a/,/i/ and /u/ contexts respectively.

(ii) +12 dB SNR presentations: With 50% FTDM, Voicing feature benefits were 0%, +11% and +10% ; while with 100% FTDM, benefits were +0%, +6% and -5% in /a/,/i/, and /u/ contexts respectively.
(iii) +6dB presentations: With 50% FTDM, Voicing feature benefits were +4%,-4%, and +8% ; while with 100% FTDM, benefits were +4%,-10% and +34% in /a/,/i/,/u/ contexts respectively.
Placing: From Table IV, (i)noise-free presentations: With 50% FTDM, Placing feature benefits were 0%, -1% and -1% ; while with 100% FTDM, benefits were 0%, 0%, and -4% in /a/,/i/ and /u/ contexts respectively.
(ii)+12 dB SNR presentations: With 50% FTDM, Placing benefits were 0%, +13% and +3% benefits; while with 100% FTDM, benefits were -2%, +1% and -3% benefits in /a/,/i/, and /u/ contexts respectively.
(iii) +6dB SNR presentations: With 50% FTDM, Placing benefits were +5%,+25%, and +37%; while with 100% FTDM, benefits were +7%,+10% and +16% in /a/,/i/,/u/ contexts respectively.

From Table IV, FTDM effects on statistical significance status are as below,
Voicing: 6dB SNR presentations, reported significant intelligibility benefits for 50% and 100% FTDM (p=0.0002 and p=0.0064) in /u/ context only; while 12dB SNR presentations failed to report any significant benefit with 50% and 100% FTDMs.
Placing: 6dB SNR presentations, reported significant intelligibility benefits for 50% FTDM (p=0.095, p=0.0003, p=0.0002) in /a/,/i/,/u/ contexts ; also 100% FTDM with 6dB SNR presentations, reported significant benefit (p=0.0131) in /i/ context only; while 12dB SNR presentations failed to report any significant benefit with 50% and 100% FTDMs.

IV. SUMMARY AND CONCLUSIONS

Two schemes of selective consonant duration modifications on fricative-vowels for speech intelligibility improvement have been proposed. Results from perception experiments suggested that, (i) ND lengthening by 50% and 100% of original duration reported decrease in performance on voicing and placing perception in the presence of noise, hence may be detrimental, (ii) FTD lengthening by 50% of original duration reported consistent significant benefit on Place of articulation feature in the presence of +6 dB noise, (iii) Both FTD and ND lengthening failed to report significant benefit on Voicing feature.

In summary, FTD has been proved as one of the acoustic cues most predictive of place of articulation decisions in the presence of noise in consistent with previous findings [13, 14, and 26]. On the other hand, ND which is the cue to fricative manner of articulation does not seem to be suitable for consonant lengthening of fricatives. The above findings suggested that efforts to emphasize potentially weak consonant should be beneficial in surmounting some of the speech recognition difficulties of hearing impaired listeners; hence should be useful in developing appropriate speech signal processing mechanisms to improve speech perception in the hard of hearing.

Table I: NDM Scheme- Perceptual Analysis

NDM Scheme – Relative Information Transmission Scores										
FEATURE	**VOWEL CONTEXT**	**NOISE FREE**			**SNR=12 dB**			**SNR=6 dB**		
		NDM(%)			**NDM (%)**			**NDM (%)**		
		0	**50**	**100**	**0**	**50**	**100**	**0**	**50**	**100**
VOICING	**/a/**	92	95	100	97	91	100	96	97	100
	/i/	94	98	90	86	90	97	97	100	90
	/u/	97	97	93	87	100	100	97	93	97
PLACING	**/a/**	100	98	98	100	100	100	97	100	100
	/i/	98	97	97	81	82	85	69	64	71
	/u/	98	98	97	91	100	98	88	87	87

Table II: NDM Scheme- Statistical Analysis

NDM Scheme – Two-tailed t-test Results																	
Feature	SNR (dB)	NDM(%)	Context-/a/					Context-/i/					Context-/u/				
			Mean	SD	Two Tailed t Test			Mean	SD	Two Tailed t Test			Mean	SD	Two Tailed t Test		
					t	p	Result			t	p	Result			t	p	Result
VOICING	Noise Free	**0**	**92**	10				**94**	11				**97**	6			
		50	**95**	6	0.514	0.6253	NS	**98**	5	0.662	0.5325	NS	**97**	5	0	1	NS
		100	**100**	0	NaN	NaN	ND	**90**	13	-0.47	0.6551	NS	**93**	9	-0.74	0.4875	NS
	12	**0**	**97**	5				**86**	28				**87**	19			
		50	**91**	18	-0.642	0.5444	NS	**90**	11	0.266	0.7992	NS	**100**	0	NaN	NaN	ND
		100	**100**	0	NaN	NaN	ND	**97**	5	0.773	0.4686	NS	**100**	0	NaN	NaN	ND
	6	**0**	**96**	9				**97**	5				**97**	5			
		50	**97**	5	0.194	0.8524	NS	**100**	0	NaN	NaN	ND	**93**	9	-0.777	0.4667	NS
		100	**100**	0	NaN	NaN	ND	**90**	20	-0.679	0.5224	NS	**97**	5	0	1	NS
PLACING	Noise Free	**0**	**100**	0				**98**	3				**98**	3			
		50	**98**	3	NaN	NaN	ND	**97**	4	-0.4	0.703	NS	**98**	3	0	1	NS
		100	**98**	3	NaN	NaN	ND	**97**	3	-0.471	0.654	NS	**97**	4	-0.4	0.703	NS
	12	**0**	**100**	0				**81**	14				**91**	11			
		50	**100**	0	NaN	NaN	ND	**82**	14	0.101	0.9228	NS	**100**	0	NaN	NaN	ND
		100	**100**	0	NaN	NaN	ND	**85**	6	0.525	0.6183	NS	**98**	3	NaN	NaN	ND
	6	**0**	**97**	4				**69**	5				**88**	7			
		50	**100**	0	NaN	NaN	ND	**64**	15	-0.632	0.5504	NS	**87**	13	-0.135	0.8967	NS
		100	**100**	0	NaN	NaN	ND	**71**	16	0.239	0.8193	NS	**87**	6	-0.217	0.8355	NS

S-Significant, NS- Not Significant, ND-Not defined

Table III: FTDM Scheme- Perceptual Analysis

FTDM Scheme – Relative Information Transmission Scores										
FEATURE	VOWEL CONTEXT	NOISE FREE			SNR=12 dB			SNR=6 dB		
		FTDM(%)			FTDM (%)			FTDM (%)		
		0	50	100	0	50	100	0	50	100
VOICING	/a/	92	97	100	97	97	97	96	100	100
	/i/	94	97	95	86	97	92	97	93	87
	/u/	97	97	90	87	97	82	63	71	97
PLACING	/a/	100	100	100	100	100	98	62	67	69
	/i/	98	97	98	81	94	82	68	93	78
	/u/	98	97	94	91	94	88	51	88	67

Table IV: FTDM Scheme- Statistical Analysis

FTDM Scheme – Two-tailed t-test Results																	
Feature	SNR (dB)	FTDM(%)	Context-/a/					Context-/I/					Context-/u/				
			Mean	SD	Two Tailed t Test			Mean	SD	Two Tailed t Test			Mean	SD	Two Tailed t Test		
					t	p	Result			t	p	Result			t	p	Result
VOICING	Noise Free	0	92	10				94	11				97	6			
		50	97	5	0.894	0.4424	NS	97	5	0.497	0.6529	NS	97	5	0	1	NS
		100	100	0	NaN	NaN	ND	95	11	0.129	0.9019	NS	90	11	-1.117	0.3581	NS
	12	0	97	5				86	28				87	19			
		50	97	5	0	1	NS	97	5	0.773	0.5077	NS	97	5	1.018	0.5174	NS
		100	97	5	0	1	NS	92	16	0.372	0.7504	NS	82	29	-0.288	0.7881	NS
	6	0	96	9				97	5				63	5			
		50	100	0	NaN	NaN	ND	93	9	-0.777	0.5127	NS	71	4	-8.121	0.0002	S
		100	100	0	NaN	NaN	ND	87	10	-1.789	0.1637	NS	97	12	-5.231	0.0064	S
PLACING	Noise Free	0	100	0	·	·		98	3				98	3			
		50	100	0	NaN	NaN	ND	97	4	-0.4	0.6923	NS	97	6	-0.298	0.7906	NS
		100	100	0	NaN	NaN	ND	98	3	0	1	NS	94	11	-0.702	0.6463	NS
	12	0	100	0				81	14				91	11			
		50	100	0	NaN	NaN	ND	94	8	1.612	0.2139	NS	94	9	0.422	0.682	NS
		100	98	3	NaN	NaN	ND	82	11	0.112	0.9119	NS	88	20	-0.263	0.8196	NS
	6	0	62	5				68	3				51	7			
		50	67	5	-1.98	0.095	S	93	5	-8.575	0.0003	S	88	6	-8.026	0.0002	S
		100	69	9	0	1	NS	78	6	-4.472	0.0131	S	67	19	-2.074	0.2713	NS

S-Significant, NS- Not Significant, ND-Not defined

REFERENCES

[1]. Krause, J.C. Braida, L.D. 2002. Investigating alternative forms of clear speech: the effects of speaking rate and speaking mode on intelligibility. J. Acoust. Soc.Am., 112, pp 2165-2172..

[2]. Picheny, M.I. Dulach, N.I. and Briada, L.D. 1985. Speaking clearly for the hard of hearing I: Intelligibility Differences between clear and conversational speech. J. speech and Hearing Res., 28, pp 96-103.

[3]. Picheny, M.I. Dulach, N.I. and Briada, L.D.1986. Speaking clearly for the hard of hearing II: Acoustic characteristics of clear and conversational speech. J. speech and Hearing Res.,. 29,pp 434-446..

[4]. Krause, J.C. Braida, L.D.2004. Acoustic properties of naturally produced clear speech at normal speaking rates. J. Acoust. Soc.Am.,. 115 (1), pp 362-378.

[5]. Shobha, N.H., Thomas, T.G., and Subbarao, K.. 2008.Influence of Consonant-Vowel Intensity Ratio on Speech Perception for Hearing Impaired Listeners, Proc. 2008 Congress in Image and Signal processing, IEEE CISP 2008, Hainan, China, Vol 5, pp 342-346, ISBN:978-0-7695-3119-9.

[6]. Shobha, N.H., Thomas, T.G., and Subbarao, K. 2010. Experimental Evaluation of Improvement in Consonant Recognition for the Hearing-impaired Listeners, IEI Journal- Interdisciplinary , INDIA ,Vol 91, pp 1-7.

[7]. Shobha, N.H., Thomas, T.G., and Subbarao, K. 2009. Effect of Consonant Duration Modifications on Speech Perception in Noise-I, International Journal of Electronics Engineering, Vol 2, Issue 2, pp 179-184.

[8]. Shobha, N.H., Thomas, T.G., and Subbarao, K. 2010. Effect of Consonant Duration Modifications on Speech Perception in Noise-II, International Journal of Electronics Engineering, vol pp 75-81.

[9]. O'Shaughnessay, D. 1981. Speech communication: Human & Machine. Massachusetts: Addison-Wesley.

[10]. Van Heuven, V.J., 'Reversal of the rise-time cue in the affricative-fricative contrast: An experiment on the silence of sound', The Psychophysics of speech perception, Martinus Nijhoff, Dordrecht, pp 181-187. 1987.

[11]. Miller G.A. and Nicely, P.E 1955. 'An analysis of perceptual confusions among some English consonants,.J. Acoust. Soc. Am., 27(2), pp 338-352.

[12]. Bilger, R.C. & Wang, M.D. (1976). Consonant confusions in patients with sensorineural hearing loss, J. Speech Hear. Res., 19, 718-748.

[13]. Jongman, A., Wayl&, R., & Wong, S. (2000). Acoustic characteristics of English fricatives, J. Accost. Soc. Am., 108, 1252-1263.

[14]. Nittrouer, S. (2002).Learning to perceive speech: How fricative perception changes, and how it stays the same. J. Accost. Soc. Am., 112, 711-719.

[15]. P Boersma, & D Weenik, PRAAT: Doing Phonetics by computer, http://www.praat.org, Version 4.4.22,2005.

[16]. Jongman, A, Spence, M., Wang, Y., Kim, B., and Schenck, D., 'Perceptual properties of English fricatives', J. Accost. Soc. Am., Vol. 105, 1401(A), 1999.

[17]. Jiang, J. Chen, M. and Alwan, A.: 'On the perception of voicing in syllable-initial plosives in noise', J. Acoust. Soc. Am., 119(2), pp 1092-1105, 2005.

[18]. Stevens, K., & Blumstein, S. 1978. Invariant cues for place of articulation in stop consonants. J. Accost. Soc. Am., 64(5), 1358-1368.

[19]. Oxenham, A.J. and Moore, B.C.J. 1987. Modeling the effects of peripheral nonlinearity in listeners with normal and impaired hearing, In W. Jesteadt (Ed.), 'Modeling sensorineural hearing loss'(Mahwah, NJ: Lawrence Earlbaum Associates) pp 273-288.

[20]. Sussman, J.E.1993.Perception of formant transition cues to place of articulation in children with language impairments. J.Speech Lang. Hear.Res., 30, pp 1286-1299.

[21]. Flecher, H..1952. The perception of sounds by deafened persons, J. Acoust. Soc. Am.., 24. Pp 490-497.

[22]. S A Simpson, & M Cooke. 2005. Consonant identification in N-talker babble in a non monotonic function of N, J. Acoust. Soc. Am.., 118(5), pp 2775-2778.

[23]. <http://www.lifesci.Sussex.ac.uk/home/Chris_Darwin/praatscripts/ >

[24]. Bilger, R.C. and Wang, M.D.1976. Consonant confusions in patients with sensorineural hearing loss. J.speech and Hearing Res., 19, pp 718-748,

[25]. Ladefoged, P.1975. A course in phonetics, Orlando: Harcourt Brace, 4th ed. 2001, Boston.

[26]. Maniwa, K., and Jongman,A.(2009). Acoustic characteristics of clearly spoken English fricatives, J. Acoust. Soc. Am.., 125(6), pp 3962-3974