

Effect of Compression Ratio and Release Time on Acoustic Cues of Speech Syllables

Hemanth N*

All India Institute of Speech and Hearing, India

***Corresponding author:** Hemanth N, Lecturer in Audiology, All India Institute of Speech and Hearing, Mysuru, India, Tel: +91 9986511550; Email: hemanthn.shetty@gmail.com

Short Communication

Volume 1 Issue 2

Received Date: August 4, 2016

Published Date: August 12, 2016

DOI: 10.23880/OOAJ-16000117

Abstract

Temporal and spectral cues play major role in speech perception. It is well documented in the literature that envelope is altered after processed through compression parameters of hearing aid. However, there is a dearth of literature on the status of inherent acoustic cues after amplification with varying compression release time and ratio. This study is investigated by recording the hearing aid output of voiced and voiceless stop consonants combined with vowel /a/ in initial and final positions. Further, a acoustic analyses such as closure duration and transition duration in temporal domain; and in spectral domain F2 transition and speed of transition were measured in different compression release times and ratios. It was noted that different compression release times and the ratios have preserved each acoustic cue of stop consonants. To conclude, irrespective of compression ratios and release times, a WDRC hearing aid does not alter the acoustic cues measured, when the signal input delivered to hearing aid was at 65 dB SPL.

Keywords: Release time; Hearing aid; Formant transition

Abbreviations: WDRC: Wide Dynamic Range Compression; EDI: Envelope Difference Index; RT: Release Time; VCV: Vowel Consonant Vowel; ET: Extent of Transition; ST: Speed of Transition Duration; CD: Closure Duration; TD: Transition Duration

Background

Hearing aid is one among rehabilitative devices that can be prescribed to any degree of hearing loss. It alleviates hearing loss by merely amplifying speech sounds above patient's threshold, so that speech spectrum will be in the audible range without any discomfort. This is done by setting the appropriate gain in each channel, based on the amount of hearing loss at each frequency. In addition, compression in the hearing aid compresses high level sounds below discomfort level and amplifies the low level sounds above the patient's

threshold. In doing so, wide range of intensities is audible to them. Thus, hearing aid acts to wide range of input signal to higher intensity as loud sound and low intensity sound as soft sounds. This wide dynamic range compression (WDRC) hearing aid closely resembles physiology of cochlea. That is, the outer hair cells in cochlea acts as a cochlear amplifier where it amplifies soft sounds, however, the inner hair cells receives direct stimulation during loud sound and further regulates in processing it by inhibitory efferent auditory nerve, thereby loud sound be well below patient discomfort level. Whereas, in a wide dynamic range compression hearing aid works linearly (constant gain) for a low level sounds (physiologically like outer hair cells) and non-linearly (variable gain) for high level sounds i.e., compression activates if the sounds are above compression threshold in hearing aid. The time taken to active compression circuit describes attack time. This is

verified by varying the input signals from low to high intensity signals. However, on the other end the time taken to release from compression is called release time. This is measured by reducing the input intensity from high to low input to hearing aid. As we know speech is dynamic and its intensity instantaneously varies its amplitude in an ongoing time. Certainly the hearing aid does react according to the input intensity signal. Now the question arises, how well the compression circuit in a hearing aid preserves the inherent acoustic cues, which helps in speech perception. In literature, it was well studied that temporal content of speech altered by compression parameters (Release time and compression ratio) in hearing aids. It was confirmed by objectively analyzing the recorded hearing aid output using envelope difference index (EDI) [1].

It was noted that sentences which were processed from shorter compression release time with higher compression ratio altered the temporal envelope, this reflects the quality and intelligibility being rated low from individuals with hearing impairment [2]. Further, some of the interesting findings realized in clinics, especially, at the time of validating the hearing aid, where majority of the time stop sounds are misperceived. In order to find out the reason behind the misperception from hearing impaired individuals, Souza, Tremblay, Davies-Venn and Kalstein [3] recorded hearing aid output in the rear ear. Acoustic analyses of unprocessed and processed (hearing aid output) of CV syllables revealed alterations in amplitude of burst spectrum. That is, the burst spectrum of /ki/ after hearing aid processing is same as that of unprocessed /ti/, because of which /ki/ was consistently misidentified as /ti/. To summarize, alteration in the envelope is more likely for a shorter compression release time and ratio, which was objectively analyzed using EDI. In addition, the shift in amplitude of the burst spectrum in aided speech leads to a misperception. In addition to the envelope and amplitude of burst spectrum, there are other cues such as closure duration, formant transition duration in temporal aspect and in spectral domain includes the extent of formant transition and speed of transition plays major cues for the perception of stop consonants. It is hypothesized that compression parameter of hearing aid might alter the inherent acoustic cues. Thus, the present study examined the effect of release time (RT) and compression ratio (CR) on acoustic cues of VCV syllables. Hearing aid having the option to change the compression release time and the ratio was utilized to record the speech sounds. A total of six stimuli, which includes three voiced and three voiceless stop consonants combined with the vowel /a/ in initial and final positions were used as target test stimuli. These

vowel consonant vowel (VCV) syllables were presented at 65 dB SPL, which was well above the compression threshold of hearing aid. The output of the hearing aid for these six target test stimuli was acoustically analyzed. The detail procedure is outlined below.

Preparation and Presentation of Stimulus

The following procedure was carried out in preparing the stimulus. The phonemes of voiced and voiceless stop consonants in Kannada were selected [/b/ & /p/ (labial); /d/ & /t/ (alveolar); /g/ & /k/ (velar)] and paired with a low short central vowel /a/ at initial and final positions. Three female speakers whose mother tongue being Kannada were chosen to utter each of these six vowel-consonant-vowel (VCV) syllables. These VCV syllables recorded using Adobe Audition (version 3) software via the recording microphone placed at a distance of 10 cm from the lips of the speaker. The recorded stimuli were digitized using a 32-bit processor at 44,100 Hz sampling frequency and finally normalized to Root Mean Square (RMS). Goodness test was performed to verify the naturality of test stimuli. A total of 18 stimuli (6 VCV stimuli * 3 speakers) were presented to ten normal hearing individuals for quality rating. Those VCV syllables which were rated more natural from one among three speakers were selected as target test stimuli. Once after selecting the target test stimuli, digitally recorded six target test VCV stimuli from same speaker were concatenated with the inter-stimulus interval of 5 sec and stored in Cubase software. The stored VCV stimuli were delivered at 65 dB SPL through a loudspeaker, positioned at one meter distance at 0° azimuth from the reference position. Reference position is a location where the MANKIN was placed.

Recording of Hearing Aid Output

Two channels behind-the-ear programmable hearing aid having the option of changing release time and the compression ratio was selected. The compression threshold of hearing aid was at 30 dB SPL. The activation of the compression circuit at a lower intensity level is referred to as wide dynamic range compression [4]. The hearing aid being programmed for calm situation with compression ratio set to 1:1 (linear). The omnidirectional microphone was enabled. The programmed hearing aid was fitted to pinna of MANKIN and output the microphone was fed into the pre-amplifier and in turn the processed speech was stored in the SLM (Bruel & Kjaer 2250) as wave format. A similar procedure was followed in setting the hearing aid at various compression release times (40 msec. 640 msec. and 1000 msec.) with

compression ratios of 2:1 and 3:1. Compression settings in hearing aid for different recording conditions are tabulated in Table 1.1. Each VCV stimulus was recorded in seven settings [(1) linear setting + (3) release times* (2) compression ratios] (Table 1). Further, the acoustic analyses were carried out on 42 VCV (7 different hearing aid setting *3 each voiced and voiceless stimuli) syllables.

Conditions	Compression ratio in both channels	Compression release time in both channels
1	1	40 msec.
2	2	40 msec.
3	3	40 msec.
4	2	640 msec.
5	3	640 msec.
6	2	1000 msec.
7	3	1000 msec.

Table 1: Compression settings of hearing aid for recording conditions.

Acoustic Boundaries

Two speech science professionals served as judge to analyze the acoustic cues of the hearing aid output recorded in different combinations of compression release time and ratio. The acoustic cues were analyzed in terms of spectral and temporal aspects. In temporal cues, closure duration and F2 transition duration were measured. The spectral cues consist of extent of transition duration and speed of transition duration. The rules used for measuring the spectral and temporal cues of voiced and voiceless VCV speech syllables are given in Table 2.

Acoustic Cues	Rule
Temporal Cues	
Closure duration (CD)	The time difference between the onset of closure and articulatory release in word medial stop consonant
F2 transition duration(TD)	Time difference between onset and steady state of second formant frequency of vowel following consonant
Spectral cues	
Extent of transition (ET)	Frequency difference between onset and steady of vowel following consonant
Speed of transition duration (ST)	Ratio of extent of transition to F2 transition duration

Table 2: Rules for measuring spectral and temporal acoustic cues in each VCV syllable.

The data of each acoustic cue from different compression release time and ratio were analyzed descriptively. From Figure 1, it was noted that for voiceless VCV syllables CD was more than voiced VCV syllables. In addition changing compression parameter has no effect on CD. In TD (Figure 1), labial voiced and voiceless VCV syllables have longer transition, followed by alveolar and velar voiced and voiceless VCV syllables. In addition, TD of each syllable was preserved in different combination of compression release time and ratio.

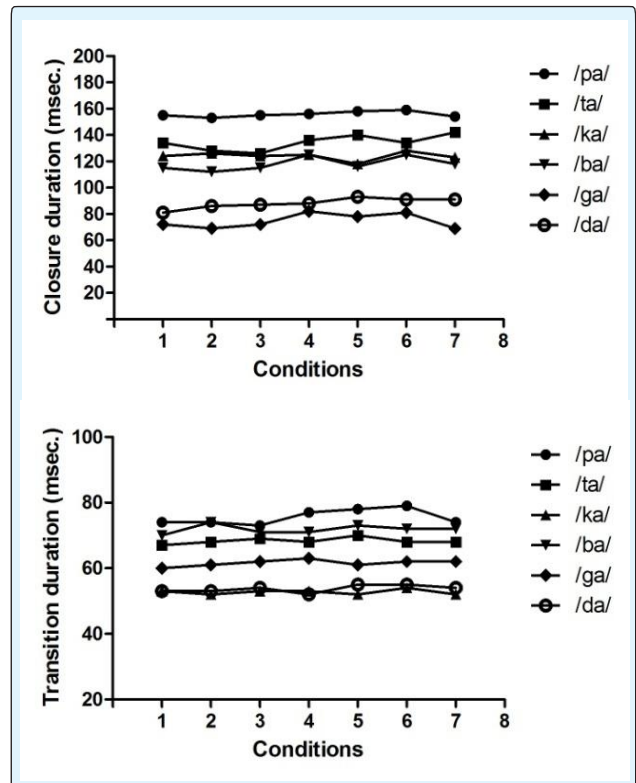


Figure 1: Closure duration and transition duration of each voiced and voiceless syllable is plotted in each condition.

In spectral parameter, formant transition and speed of transition was measured. From Figure 2 raising pattern of formant transition was noted in both voiced and voiceless labial VCV syllables. However, in alveolar and velar voiced and voiceless VCV syllables, falling pattern of transition was noted. It was noted that F2 transition in each syllable was preserved with varied compression release time and ratio. Further, speed of the transition from (Figure 2) is faster for velar than alveolar VCV syllable followed by the labial VCV syllable. This is true for both voiced and voiceless VCV syllables. Speed of transition was preserved with different compression release time and ratio. To conclude, different combination compression release time

and ratio have preserved each acoustic cue which is important for the perception of voiced and voiceless stop consonants. Further, research has to be carried out by varying the input intensities and see to it is that hearing aid preserves inherent acoustic cues.

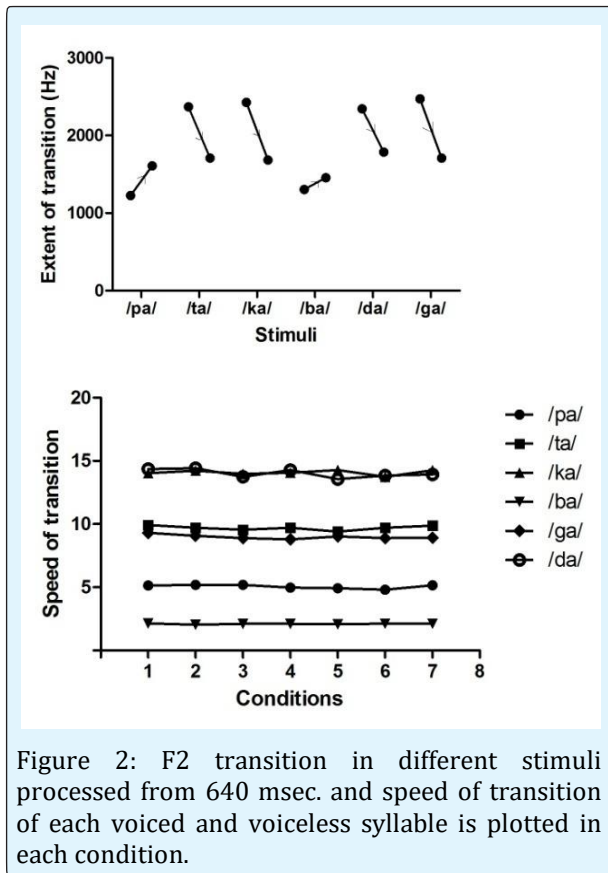


Figure 2: F2 transition in different stimuli processed from 640 msec. and speed of transition of each voiced and voiceless syllable is plotted in each condition.

Acknowledgement

The authors would like to thank the Director, All India Institute of Speech and Hearing for granting permission to carry out the study; the authors would also like to thank HOD, Department of audiology, for permitting us to use the instruments necessary to carry out the study.

References

- Hickson L, Thyer N (2003) Acoustic analysis of speech through a hearing aid: perceptual effects of changes with two-channel compression. *J Am Acad Audiol* 14(8): 414-426.
- Jenstad LM, Souza PE (2005) Quantifying the effect of compression hearing aid release time on speech acoustics and intelligibility. *J speech Lang Hear Res* 48 (3): 651-667.
- Souza P, Tremblay K, Davies-Venn E, Kalstein L (2004) Explaining consonant errors using short-term audibility. Presented at: the American Academy of Audiology Salt Lake City Utah.
- Dillon H (2001) *Hearing Aids*. New York NY: Thieme Medical Publishers.

