

Low Frequency Perception of Rhythm and Intonation Speech Patterns by Normal Hearing Adults

Youngsun Kim* · Carl W. Asp*

ABSTRACT

This study tested normal hearing adults' auditory perception of rhythm and intonation patterns, with low-frequency speech energy. The results showed that the narrow-band low-frequency zones of 125, 250, or 500 Hz provided the same important rhythm and intonation cues as did the wide-band condition. This suggested that an auditory training strategy that uses low-frequency filters would be effective for structuring or re-structuring the perception of rhythm and intonation patterns. These filters force the client to focus on these patterns, because the speech intelligibility is drastically reduced. This strategy can be used with both normal-hearing and hearing impaired children and adults with poor listening skills, and possibly poor speech intelligibility.

Keywords: Rhythm, Intonation, Low Frequencies, Perception, Auditory Training

1. Introduction

As a part of normal development, an infant babbles and engages in a pre-word dialogue with his mother or caregivers (Crystal, 1987). This dialogue begins with a strong emotional bond between the infant and mother, where the infant can both feel and hear the mother's emotional vocal patterns. This pre-word dialogue has both normal rhythm and intonation speech patterns that are meaningful and critical for the infant's development (Asp, 2002). The early development of these patterns provides the foundation both for spoken language and listening skill, even though there are numerous speech errors of undeveloped phonemes.

As young child advances to the word level, his speech becomes more intelligible. By eight years of age, the child mastered the accuracy of all the 42 phonemes of English, and is completely intelligible in a speech dialogue with others. However, the foundation for this dialogue is the normal rhythm and intonation patterns that were developed as an infant (Asp, 2002).

* Dept. of Audiology and Speech Pathology, University of Tennessee, Knoxville

Other investigators (Lehiste, 1976; Crystal, 1987) have studied the importance of rhythm and intonation patterns by identifying them as the suprasegmental or prosodic features of spoken language. Hargrove and McGarr (1994) describe speech rhythm as a facilitator of monitoring content of the speaker's message and improving the speech intelligibility. To communicate effectively, all normal listeners anticipate speech rhythms in a dialogue. This anticipation of the rhythm patterns makes it possible for the listeners to understand spoken language, and monitor their own speech. However, when speakers violate the listeners' rhythm expectations, the listeners become attentive to the linguistic content. In addition, the anticipation of rhythmic patterns helps the speaker use the proper timing of rhythmic patterns (Asp, 2002).

Kent and Read (1992) described the intonation speech patterns as providing both the emotional content and the grammatical structure that make speech meaningful. A rising intonation is used for a question, whereas a falling intonation is for a statement. Brazil et al. (1980) indicated that the speaker's intonation patterns convey the speaker's attitude and social status. Asp (2002) described normal intonation patterns as being necessary for developing good social skills and high emotional quotient (EQ). However, the importance of rhythm and intonation patterns has been neglected in the routine investigation for speech and hearing impairment, even though the suprasegmentals are the basis of good spoken language (Lehiste, 1976).

How can rhythm and intonation patterns be used to help children and adults with communication disorders and differences? For the client with hearing impairments, the investigators agree that the residual hearing is essential for developing good spoken language and listening skill. In 1980, Guberina and Asp identified the low-frequency zone below 500 Hz because it has more hearing sensitivity for the hearing impaired, and the rhythm and intonation patterns can be processed through this zone. This strategy was to use an auditory training unit with an extended low-frequency range below 125 Hz. This unit included a vibrotactile input to feel the speech rhythms, a headset to hear them, and an acoustic filter to maximize the client's low-frequency perception. As the low-frequency perception developed, the clients were able to hear phonemes in the high-frequency zone. The key was maximizing the low-frequency zone.

Ling (1964) showed that a flat extended low-frequency response down to 100 Hz resulted in higher speech intelligibility in children. Rosenthal, et al. (1975) reported that low-frequency speech energy in combination with high-frequency energy increased the correct identification of consonants by more than 20%. Rhodes (1966) reported that hearing-impaired clients with good low-frequency sensitivity below 1kHz had significantly better listening scores than normal-hearing people, when the speech

signal was passed through a 1,000 Hz low-pass filter. Rhodes explained that hearing-impaired listeners apparently use acoustic cues not normally recognized by normal-hearing listeners. This suggests that low-frequency residual hearing has potential to improve oral communication skills.

In contrast to the above, the famous Harvard Report (1947) recommended that the frequency response of hearing aids for adults should not amplify the area below 300 Hz, because of the upward spread of masking from low-frequency ambient noise. This report created a negative impression of the low-frequency speech energy zone for both auditory training and for hearing aid placement. In addition, the recommendation was applied to young children even though no children were involved in the research project.

In search of better understanding of the low-frequency zone, and rhythm and intonation patterns, the basic research question was; Does the low-frequency zone provide speech cues for auditory perception of rhythm and intonation patterns? To answer this question, the current study separated the low-frequency speech zone with 125, 250, and 500 Hz narrow-band low-pass filter and compared the test results to a standard wide-band reference condition. The research question was; How does the low-frequency zone compare to the wide-band zone for the auditory perception of rhythm and intonation patterns, using normal-hearing adult subjects.

2. Methods

2.1 Stimuli and Recording Procedure

The test stimuli consisted of six rhythm and six intonation speech patterns (see Figure 1). Each pattern had the nonsense syllable /bʌ/ repeated four times. The rhythm patterns had both stressed (●) and unstressed syllables (○), with either regular or quick tempo. The quick tempo had two-syllables close together, e.g., ○○, whereas the regular tempo had the syllables separated in regular intervals, e.g., ○○. The intonation patterns consisted of the /bʌ/ syllables, with each syllable having a rising or a falling intonation pattern, e.g., ↘ ↘ ↗ ↗.

An experienced clinician in rhythm and intonation therapy vocally produced two separate randomized test lists using the nonsense syllable /bʌ/. Each list was recorded on a separate audio cassette tape, using a tape recorder (Marantz, model PMD 430) in a quiet sound treated booth. Two experienced clinicians independently judged the two lists to have similar patterns. The mean duration of rhythm patterns was 1.25 seconds with a range of 0.98 to 1.42 sec, whereas the mean for the intonation patterns was 3.83 seconds, with a range of 3.5 to 4.9 sec.

2.2 Subjects and Procedures

Twelve normal-hearing adults listened and imitated each of the six rhythm and the six intonation speech patterns, under four different test conditions. These conditions include a wide-band (20 to 20,000 Hz), and three low-pass filter conditions of 125, 250 and 500 Hz, with a sharp slopes of 60 dB per octave (see Figure 2). The sharp slope attenuated speech energy above the cutoff frequency. These cutoff frequencies corresponded to the three pure-tone frequencies used in standard audiometric testing. All four test conditions were easily set on a high-quality auditory training unit (Listen II, model 1000) and played through a high-quality loudspeaker (JBL model pro III) in a sound treated booth. With a listener seated at a 3-foot distance, the wide-band condition was the first listening condition; it was followed by the three low-pass filter conditions in a randomized order. To minimize learning effect, the two test lists were used to create a different randomized order for each subject.

While listening through the 250 Hz low-pass condition, the experimenter and an audiologist independently set the training unit amplifier at their Most Comfortable Loudness Level (MCL); this setting was measured at 86 dB SPL. With the amplifier at this setting, the other low-pass speech levels were measured at 74 dB SPL for 125 Hz, and 92 dB SPL for 500 Hz. For the wide-band condition, the experimenter's and the audiologist's MCL was 70 dB SPL. With a similar MCL for all four conditions, the difference in SPL was a result of the speech spectrum through the four different bandwidths. In comparison, Rosenthal et al.(1975) recommended that narrow-band SPL levels be set 10 to 30 dB above the wide-band normal speech conversation level (65 dB SPL), because narrow-bands need more SPL than wide-bands to achieve the same loudness.

All of the subjects passed a training session in the wide-band condition by vocally imitating at least 90% of 10 practice items correctly. This verified that each subject had the skills and understanding to complete the listening tasks. For the experiment, each vocal imitation of each subject was recorded on an audio tape. The experimenter transcribed and judged each response as correct or incorrect. A correct response was judged to be identical to the rhythm and intonation speech pattern of the test items.

To verify the experimenter's judgement, a second judge, who was experienced, certified audiologist judged the same subject. The inter-judge reliability of two judges was $r=0.99$. In addition, the intra-judge reliability was $r=0.91$ for the experimenter re-testing one subject two weeks after the initial testing. Both the inter- and intra-judge reliability were considered satisfactory for this experiment.

| Rhythm Patterns | | | | Intonation Patterns | | | | |
|--|----|----|----|--|---|---|---|---|
| 1. | ○ | ● | ○○ | 1. | ↘ | ↗ | ↘ | ↗ |
| 2. | ● | ○○ | ● | 2. | ↗ | ↘ | ↘ | ↗ |
| 3. | ○ | ● | ○ | 3. | ↗ | ↗ | ↘ | ↘ |
| 4. | ○ | ● | ● | 4. | ↘ | ↗ | ↗ | ↘ |
| 5. | ● | ○ | ● | 5. | ↘ | ↘ | ↗ | ↗ |
| 6. | ○○ | ● | ○ | 6. | ↗ | ↘ | ↗ | ↘ |
| ●: stressed, ○: unstressed, ○: regular, ○○: quick | | | | ↗: rising intonation, ↘: falling intonation | | | | |

Figure 1. Rhythm and intonation speech patterns using the nonsense syllable /bʌ/

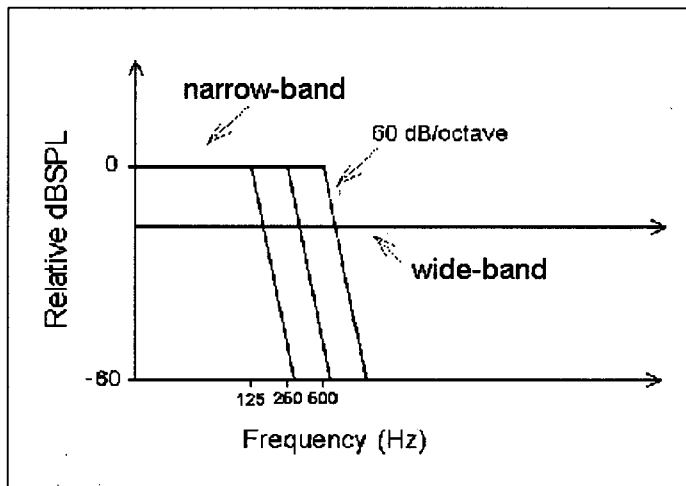


Figure 2. Four test conditions included 125, 250, 500 low-pass filters, with 60 dB slope and a wide-band condition

3. Results

A Randomized Complete Block (RCB) design, with subjects as blocks was used to compare the mean percent of correct rhythm and intonation test items across the four test conditions. The four conditions were analyzed separately for the rhythm and for the intonation patterns; the significant level was set at $p < 0.05$.

3.1 Rhythm Patterns

For the rhythm patterns, the mean percent correct had a one percent difference (99 - 100%) across the four test conditions (see Table 1 & Figure 2); there was no significant difference ($F=2.2, p=0.1$). A Post-Hoc Tukey analysis showed that three narrow-band conditions (125, 250, and 500 Hz) were not different from the wide-band

reference condition (Table 2). All three low-pass filter conditions provided the same rhythmic information as did the wide-band condition. It appeared that the bandwidth did not affect the perception of the rhythm patterns.

There were only two errors for all four conditions. One error was the omission of an unstressed syllable (● ○ ● ___ for ● ○ ● ○), while the other error was a substitution of a stressed syllable for an unstressed syllable (○ ● ○ ● for ○ ● ○○). This substitution error incurred a tempo error as a regular tempo substituted for a quick tempo. However, the perception of speech rhythm patterns had a high level of accuracy for all the subjects in both the narrow-bands and the wide-band.

3.2 Intonation Patterns

For the intonation patterns, the mean percent correct was 100%, 99.3 %, 97.2 % and 94.4%, respectively, for the four conditions. These mean scores gradually decreased (100% to 94%) as the bandwidth became narrower (see Table 1 & Figure 2). A Post-Hoc Tukey analysis showed that only the 125 Hz narrow-band condition was significantly lower than the wide-band condition (see Table 2); the mean difference was 5.6% (100 vs. 94.4%).

Table 1. The Mean and Standard Deviation (SD) of the Percent Correct (%) for the Rhythm and Intonation Patterns

| Conditions | Rhythm (N=12) | | Intonation(N=12) | |
|------------------------|---------------|-----|------------------|-----|
| | Mean | SD | Mean | SD |
| Wide-band condition | 100 | 0 | 100 | 0 |
| 500 Hz low-pass filter | 100 | 0 | 99.3 | 1.6 |
| 250 Hz low-pass filter | 100 | 0 | 97.2 | 5.2 |
| 125 Hz low-pass filter | 99.3 | 1.6 | 94.4 | 6.2 |

For the 125 Hz test condition, five of the twelve subjects (42%) perceived all the intonation patterns correctly (100%). However, seven subjects (58%) had percent correct scores ranging from 85% to 98%. Three of seven subjects substituted a ↘ ↗ ↘ ↗ pattern for ↘ ↘ ↗ ↗ pattern. Overall, subjects made nine substitutions for rising intonation and seven substitutions for falling intonation. The rising intonations had more errors than the falling intonations. However, the overall 94.4% percent correct for all subjects was still high level of performance in the narrowest bandwidth of 125 Hz.

Table 2. Post-Hoc analysis (Tukey) for paired comparison

| Filter Comparison | Rhythm | Intonation |
|-------------------------------|--------|------------|
| Wide-band vs. 500 Hz low-pass | 1.00 | .97 |
| Wide-band vs. 250 Hz low-pass | 1.00 | .36 |
| Wide-band vs. 125 Hz low-pass | .18 | .01* |

* Significant at 0.05 level

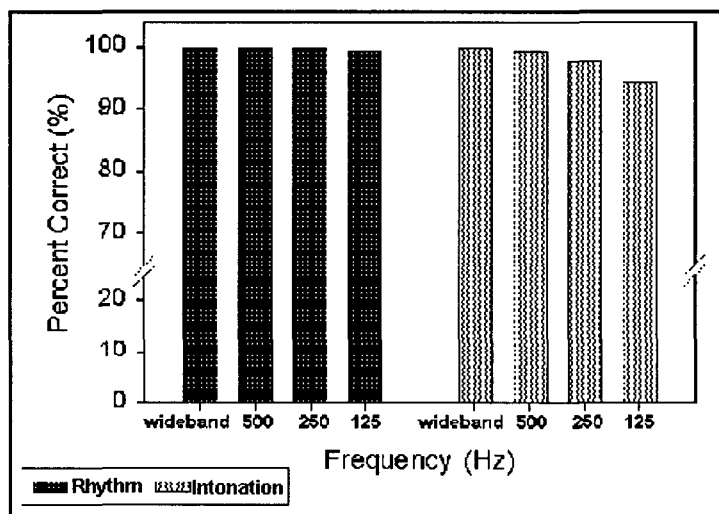


Figure 3. Mean percent correct (%) of rhythm and intonation patterns for a wide-band and three low-pass conditions of 125, 250, and 500 Hz

4. Discussion

The test result showed that the low-frequency zone provides similar speech cues for perceiving both the rhythm and intonation patterns correctly. This low-pass filters forces the listeners' focus on the rhythm and intonation patterns, because the speech intelligibility is drastically reduced. Therefore, auditory training with low-pass filters helps the listeners restructure his rhythm and intonation skills; this skills provide the foundation for the perception of segmental phonemes, grammatical structure, and emotional vocal patterns of speaker. (Asp, 2002; Kent and Read, 1992).

As mentioned earlier, some subjects (42.5%) were more skilled at perceiving the intonation patterns than others (58%) in the 125 Hz condition. Fastl and Stoll (1979) attributed this variability in the filter condition to the weakness of the pitch strength

of the intonation patterns. Therefore, an effective intonation training program, with low-pass filters, improves the listener's perception and minimize the variability among subjects.

Problems in rhythm and intonation perception occur in normal-hearing children with Central Auditory Processing Disorders (CAPD) and/or Learning Disability (LD). For example, in a case study, Earl et al. (1991) reported that a child with CAPD was initially unable to perceive and produce rhythm and intonation patterns. However, after intensive rhythm and intonation auditory training with low-frequency speech energy, the child showed a significant improvement in speech perception, auditory memory, and reading comprehension. In a follow-up group study, Earl and Rook (1994) continued the low-frequency rhythm and intonation auditory training for nine children and one adult, all of which, had CAPD. All of the nine children and the adult showed significant improvement in both auditory comprehension and auditory memory skills. The parents reported improved academic performance in reading, spelling, and handwriting. In addition, the adult was more efficient in communication skills at work after the training.

In similar study, Hall (1995) reported that the rhythm perception of 3rd graders with learning disability was 22% lower than the children with normal learning ability. He recommended auditory training of the rhythm patterns to improve the listening and memory skills; this improvement has a positive effect on the academic performance of the children.

Rhythm and intonation problems are common for hearing impaired children and for some adults. They all usually have residual hearing in the low frequency zone at below 500 Hz. These make them a good candidate for rhythm and intonation auditory training. After the training, Williams (1976) reported 56% correct for 3 to 8 years old, Strusinski (1996) reported 93% correct for 6 to 12 years old, and Asp, et al. (1990) reported 95% correct for 14 to 18 years old. All of these hearing impaired children had auditory training that emphasized low-frequency speech energy through a training unit. This training made their residual hearing functional, which in turn, improved their listening and spoken language skills.

Recently, Asp (2002) described an Auditory-Vestibular Treatment Protocol for children and adults with communication disorders and differences. This strategy emphasizes using the low-frequency zone to improve both spoken language and listening skills of children and adults. Since, both rhythm and intonation can be perceived in the low-frequency zone, the prognosis is good, if an effective treatment protocol is used. On a regular basis, the Verbotonal Speech-Sciences Research Laboratory at University of Tennessee is applying this protocol to auditory related

disorders and differences to determine the efficacy of this strategy.

References

- Asp, C. 2002. *Feel the Movement and Hear the Speech*, Knoxville: Listen.
- Asp, C., M. Kline, P. G. Duff & K. Davis. 1990. "Verbotonal method integrated into hearing services of Knox County School System." *SUVAG*, 3, 1-2.
- Brazil, D., M. Coulthrad & C. Johns. 1980. *Discourse intonation and language teaching*. Essex, England: Longman.
- Crystal, D. 1973. "Non-segmental phonology in language acquisition: a review of the issues." *Lingua*, 32, 1-45.
- Crystal, D. 1979. "Prosodic development." In P. Fletcher & M. Garman (Eds.), *Language Acquisition* (pp. 33-48). Cambridge: Cambridge University Press.
- Crystal, D. 1982. *Profiling Language Disability*. London: Edward Arnold.
- Davis, H., S. S. Stevens, R. H. Nichols Jr., C. V. Hudgins, G. E. Marquis, G. E. Peterson & D. A. Ross. 1947. *Hearing aids, An experimental study of design objectives*. Harvard, Cambridge, MA.
- Earl, D. C., C. W. Asp & L. H. Rook. 1991. "Auditory processing disorder and verbotonal listening therapy: a case study." *Poster presented at American Speech and Hearing Association Convention*. Atlanta, Georgia.
- Earl, D. C. & L. H. Rook. 1994. "Suprasegmental multisensory aural remediation techniques for treatment of auditory processing disorders." *Poster presented at American Academy of Audiology Convention*. Richmond, Virginia.
- Fastl, H. & G. Stoll. 1979. "Scaling of pitch strength." *Hearing Research*, 1, 294-301.
- Grant, K. W. & B. E. Walden. 1996. "Spectral distribution of prosodic information." *Journal of Speech and Hearing Research*, 39, 228-238.
- Guberina, P. 1964. *Verbotonal method and its application to the rehabilitation of the deaf*. Zagreb, Croatia.
- Guberina, P. & C. W. Asp. 1980. *The Verbotonal method for rehabilitating people with communication problems*. New York, NY: World Rehabilitation Fund.
- Hall, F. 1995. *Comparison of the listening skills of third-grade children with learning disabilities and normal learning abilities for word identification in noise, direct recall, and imitation of syllable rhythm patterns*. Unpublished Doctorate Dissertation, University of Tennessee, Knoxville.
- Hargrove, P. M. & N. S. McGarr. (Eds.). 1994. *Prosody management of communication disorders*. San Diego: Singular Publishing Group.
- Kent, R. & C. Read. 1992. *The acoustic analysis of speech*. San Diego: Singular Publishing Group.
- Kolike, K. & C. W. Asp. 1981. "Tennessee test of rhythm and intonation." *Journal of Speech and Hearing Disorders*, 46, 81-87.
- Ling, D. 1964. "Implication of hearing aid amplification below 300 cps." *Volta Review*,

66m 723-729.

- Leshiste, I. 1976. "Suprasegmental features of speech." In N. J. Lass (Ed.), *Contemporary issues in experimental phonetics* (pp. 225-239). New York: Academic Press.
- Rhodes, R. C. 1966. "Discrimination of filtered CNC lists by normal and hypacusics." *Journal of Auditory Research*, 6, 129-133.
- Rosental, R. D., J. K. Lang & H. Levitt. 1975. "Speech reception with low-frequency speech energy." *Journal of Acoustical Society of America*, 57, 949-955.
- Strusinski, M. 1996. *Evaluation of Verbotonal program*. Dade County Public Schools, Miami, Florida.
- Williamson, V. 1978. *Suprasegmental skills, segmental skills, and word intelligibility of hearing impaired and normal hearing children*. Unpublished master's thesis, University of Tennessee, Knoxville.

Received: Jan. 25, 2002.

Accepted: Mar. 8, 2002.

▲ Youngsun Kim
Department of Audiology and Speech Pathology
University of Tennessee
Knoxville, Tennessee 37996-0740
USA
Tel: +1-865-974-4775 Fax: +1-865-974-1539
E-mail: ykim5@utk.edu

▲ Carl W. Asp
Department of Audiology and Speech Pathology
University of Tennessee
Knoxville, Tennessee 37996-0740
USA
Tel: +1-865-974-4775 Fax: +1-865-974-1539