PROCEEDINGS of the
23rd International Congress on Acoustics

9 to 13 September 2019 in Aachen, Germany

The detection of the first and last pulses in a periodic pulse train: the role of temporal integration and masking

Liudmila RIMSKAYA-KORSAKOVA[1]; Dmitry NECHAEV[2]

[1] N.N. Andreyev Acoustics Institute, Russia

[2] A.N. Severtsov Institute of Ecology and Evolution, Russia

ABSTRACT

A pulse train containing 12 pulses with inter-pulse intervals of 20 to 110 ms was considered a model of the test and masking phonemes. The first or last pulse of the train varied in level; they were considered tests to measure the threshold difference between the test pulse alone and the same pulse in the train (dIso) and the minimum inter-pulse interval (Tmin) for the test pulse detection. The lowest Tmin and dIso were found in normal-hearing listeners. Higher dIso were found in listeners with prolonged temporal integration of periodic sounds independently of hearing sensitivity. Increased dIso (up to approximately 20 dB) were observed at Tmin despite of the test pulse position in the pulse train. The detection properties of the first and last pulses were supposed to be closely related to the well-known effects of partial masking of loudness and forward masking. Thus, within-train pulses affected the detectability of the first and last pulses. It is suggested that this effect is one of causes of changes in phonemes loudness and the speech intelligibility reduction.

Keywords: partial masking, forward masking, single pulse, pulse train, phonemes loudness, speech intelligibility

1. INTRODUCTION

The increasing of speech rate may result in reduction of speech intelligibility. This effect can be related with auditory temporal processing, but not with auditory sensitivity (1,2).

A speech can be considered a sequence of phonemes. It has been supposed (3) that an increase of the speech rate reduces the understanding of phonemes in a stream of other phonemes, and impairment of temporal discrimination of phonemes loudness can be a reason of the reduction speech intelligibility.

The consonant phonemes play a general role in speech intelligibility. The most of them have the frequency range from 1 to 5 kHz, short duration (10-70 ms) and rather low intensity. A speaker can pronounce 9-14 phonemes per second in normal speech rate which corresponds to phoneme period of 110-70 ms.

The high frequency pulse can be used as a model of consonant and the pulse train as a model of a sequence of consonants (1,2). Estimate of loudness detection of a single pulse in a pulse train with may help to understand the role of consonants loudness in intelligibility of the fast speech. For this purpose, we investigated conditions where loudness is a cue for detection of a single pulse in a train. We estimated the dependence of loudness detection of the first and last pulses on the inter-pulse interval in the train. We looked at indicators that could be characteristics of the auditory temporal processing of sounds; including speech. The loudness of the sound signal is related to auditory temporal integration process. Because of this the indicators were matched with temporal integration properties for tones and pulse train as well as with detection threshold of single pulse.

We expected that changes in the loudness of the first and last pulses would be associated with the known effects of both partial masking of loudness (4) and forward masking (5).

[1] lkrk@mail.ru
[2] dm.nechaev@yandex.ru

2. PSYCHOACOUSTIC MEASUREMENT

2.1 Methods

The psychoacoustical experiments were carried out in a sound attenuating booth. The hardware-software complex was used for generation sounds, experimental control and collection data.

One probe contained two sounds (reference and target) with 500 ms pause. The sounds were generated digitally online and played back in random order. The probes were digital/analog converted using E-MU 0204 USB with sampling rate 44.1 kHz. The analog signals were played monaurally through HD-265 headphones (Sennheiser, Wedemark, Germany).

The temporal profiles of all pulses were identical. The central frequency of the pulses was 4 kHz and bandwidth was 240 Hz (less than the critical band). The reference stimuli contained 11 identical pulses. The target stimuli contained 11 reference pulses plus one test pulse. All listeners couldn't discriminate the durations of reference and target stimuli containing 11 or more pulses. Thus, only the pulse loudness could be a cue for detection of the test pulse.

The two experimental paradigms were used: with the test pulse was either before or after pulse train (Fig. 1).

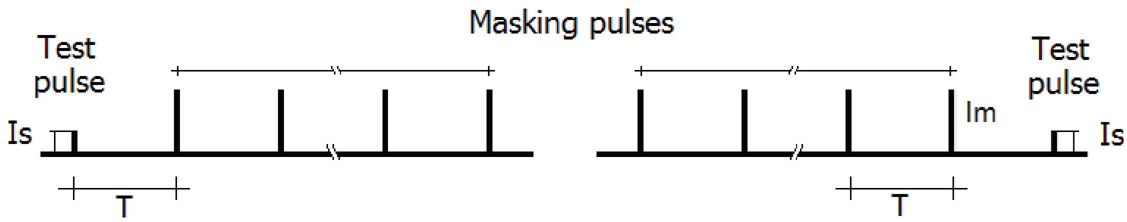

Figure 1. – Temporal diagrams with the test pulse before the pulse train (left) and after the pulse train (right). Im - amplitude of the train pulses, Is – amplitude of test pulse

The level of pulses in the train was 80 dB SPL. The level of the test pulses varied. The inter-pulse interval T in the pulse train varied from 20 to 110 ms.

The method of constant stimuli was used to obtain the psychophysical curves or the dependence of correct detection of test pulse on its intensity. Each level of the test pulse was presented 10 times, the levels alternating randomly. The level of the test pulse that provided a probability of 75% correct detection was assessed the threshold. We measured two parameters: the minimal inter-pulse interval (Tmin) in which the listeners detected the test pulse, and the difference (dIso) between the test pulse detection threshold and detection threshold of single pulse (dIso=MinIs-Min0Is, where Min0Is – the detection threshold of single pulse, MinIs – the detection threshold of the test pulse in pulse train).

We assumed that the indicator Tmin can characterize the temporal resolution of test pulse loudness in the train; the indicator dIso – changes in the loudness of the test pulses when inter-pulse interval T was near Tmin.

The indicators Tmin and dIso were compared with the temporal integration of tones and a sequence of pulses, as well as with the detection threshold of single pulses. A two-alternative forced-choice procedure combined with adaptive "two-down, one-up" stimuli trial-by-trial variation was used for determining the detection thresholds of single pulses, tones and pulse trains.

The difference between detection thresholds for tones of the some duration and tone with duration of 240 ms was used as indicators of temporal integration for tones (I_t). The tone duration was varied from 15 ms to 240 ms.

The difference between the detection thresholds for the single pulse and the pulse trains was used as indicators of train temporal integration for periodic sounds (I_ps). The inter-pulse interval in the train was 20, 30, 50, or 80 ms, therefore the train duration with 11 pulses was 200, 300, 500, or 800 ms.

2.2 Listeners

Four listeners with normal hearing (L1, L5, L6, L7) and one listener with aging-related hearing loss (L2) participated in the study. The information about listeners (age, gender, experience in psychophysical experiments, and the detection threshold of single test pulse) is presented in table 1. The results of L1, L6, L7 were averaged and denoted as typical listener (Lt), L5 was denoted as untypical.

Table 1 – information about listeners

Listener	Age	Gender	Threshold	experience; *- amateur musician
L1	35	M	17 dB	>60 hour *
L5	35	M	16 dB	Without experience *
L6	25	F	10 dB	>20 hour
L7	35	M	15 dB	<20 hour *
L2	65	F	26 dB	>60 hour

2.3 Results

Psychophysical curves which are dependences of test pulse detection on relation between levels of test pulse, Is, and intensity of pulse train, Im, are presented in Fig. 2.

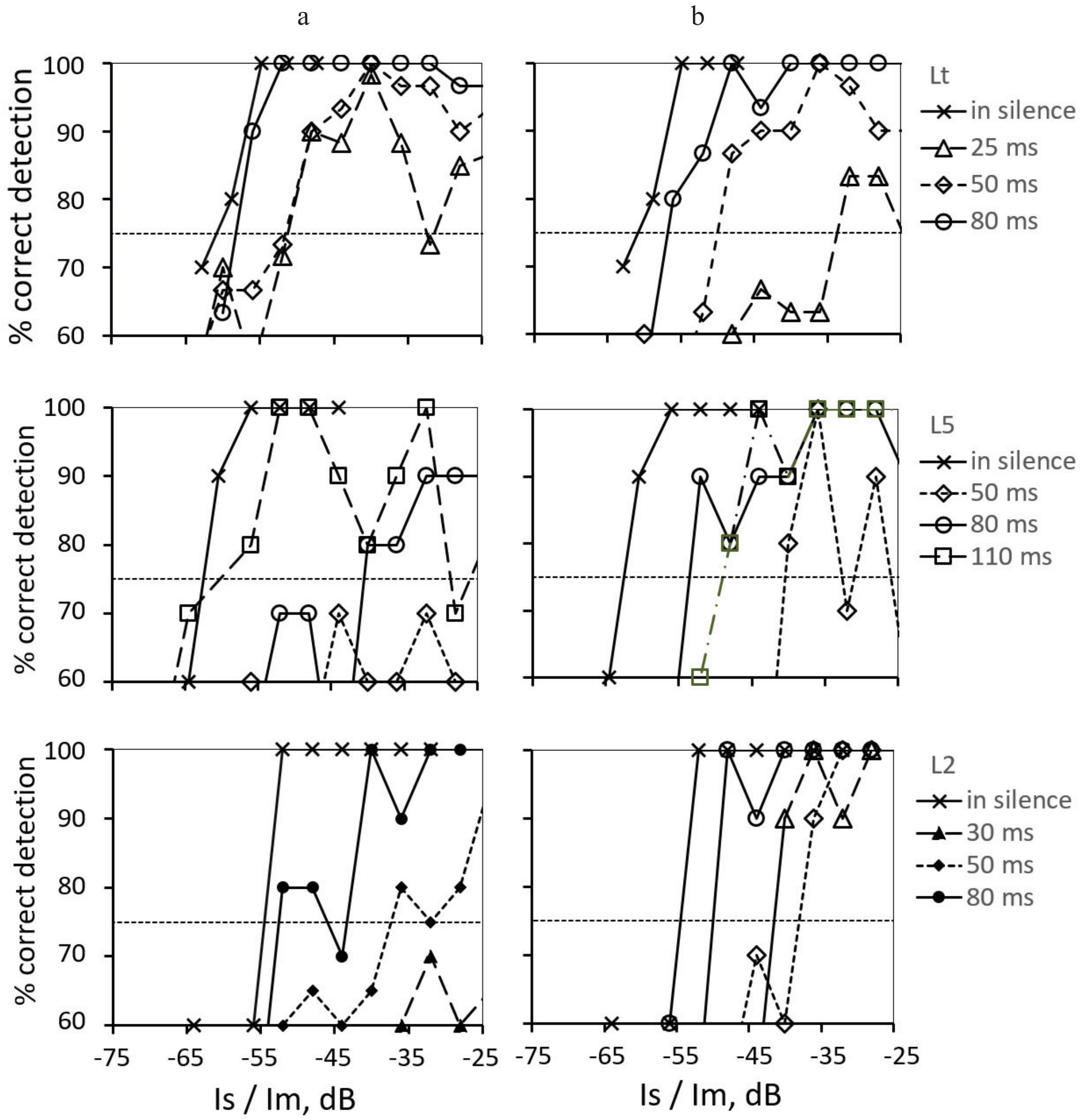

Figure 2. – Dependent of correct detection of test pulse on ratio of the test pulse level to pulse train level. (a) test pulse before the pulse train (b) test pulse after the pulse train. Data for “typical” listeners with normal hearing (Lt), “untypical” listener with normal hearing (L5) and listener with aging-related hearing loss (L2). Legend: detection test pulse without pulse train (silent) and different values of T

The values of Tmin were different for different listeners and depended on position of the test pulse in the pulse train. The lowest Tmin (25ms) was found for the typical listeners (Lt) at all position of the test pulse. The listeners L5 and L2 had different thresholds depending on the test pulse position. If the test pulse was before the pulse train, Tmin was 80 and 50 ms, resp. If the test pulse was after pulse train, Tmin was 50 and 30 ms, resp. However, the detection thresholds for single pulse were similar for listeners Lt and L5 and lower for listeners L2. Thus, the temporal resolution threshold, Tmin, for the test pulse loudness was not associated with the detection thresholds for single pulse.

The maximal values of the dIso were found for T equal to Tmin. Increasing T over Tmin caused decreasing dIso for all listeners. We suppose that the dIso could be related to changes of test pulse loudness.

For listener Lt, the highest dIso was 27 dB at T=25 ms, when the test pulse was after pulse train. When the test pulse was before pulse train, dIso was 9 dB.

The listener L5 had dIso about 23 dB at T=50 ms, when the test pulse was after pulse train and 23 at T=80 ms when the test pulse was before pulse train.

The listeners L2 had dIso about 14 dB at T=30 ms, when the test pulse was after pulse train, and 17 dB at T=50 ms, when test pulse was before pulse train.

The values of dIso and Tmin were not related to the auditory sensitivity of the listeners. The highest Tmin and dIso were found in "untypical" listener L5. The lowest Tmin and dIso was found in normal-hearing listeners.

Table 2. – The values of threshold shifts dIso

Listener	T, ms	Start of sequence dIso, dB	End of sequence dIso, dB
Lt	25	9	27
	50	9	9
	80	3	4
L5	50	-	23
	80	23	8
	110	2	14
L2	30	-	14
	50	17	16
	80	2	6

The next stage of the study was the comparison of values Tmin and dIso with the indicators of temporal integration for tones and trains. The detection thresholds of different duration tones and 11 pulse trains were presented in tables 3 and 4.

The indicators of I_t were similar for Lt and for L5 and lower for L2. The listener with aging-related hearing loss had decreased of temporal integration.

Table 3. – Indicators of temporal integration for tone

Listeners	Threshold for tone with duration 240 ms, dB	Tone duration (D_t), ms			
		15	30	60	120
		Indicators of temporal integration for tones, I_t, dB			
Lt	3,9	11,3	7,4	4,9	1,5
L5	0,9	10,9	8,0	5,0	3,0
L2	15,0	7,4	5,9	2,6	0,7

The indicator of I_ps was approximately matched the detection thresholds of single pulse for Lt. Thise indicator were varied within 2 dB for T from 20 to 80 ms. This may identify to weak summation of pulse train.

The indicator of I_ps decreased by 5-6 dB with increasing inter-pulse interval from 20 to 80 ms for listener L5. However, for the listeners L2, the indicator of I_ps was 6 – 8 dB at all T, finding remarkable summation of pulse train.

Table 4. – Indicators of temporal integration for pulse train.

Listeners	Pulse detection threshold, dB	Inter-pulse interval for 11 pulse train, ms			
		20	30	50	80
		Indicator of temporal integration for trains, I_sp, dB			
Lt	14,0	0,9	1,8	0,3	-0,6
L5	16,0	9,5	9,1	6,9	4,9
L2	26,0	7,7	6,4	8,5	7,9

Temporal integration for tones did not correspond to temporal integration for trains. The listeners Lt featured the decrease in temporal integration for train in comparison with the listeners L5 and L2. This suggests that the reason of Tmin and dIso inter-individual difference is the difference in temporal integration of pulse trains.

Integration of pulse train with inter-pulse intervals of 20-80 ms and train durations of 200-800 ms was obtained for listeners L5 and L2, but integration of tones with the same duration was absent.

It can be assumed that the integration of train, but not the integration of tones, determines the temporal resolution of the test pulse loudness.

Thus, the positive values of dIso were found in all position of pulse at T near Tmin. The lower Tmin was found in listeners with normal hearing. The high Tmin was found for listeners with different auditory sensitivity but with temporal integration for pulse train having inter-pulse interval of 20-80 ms.

3. Discussion and conclusions

In our experiments, listeners could detect the test pulse only basing on its loudness. We suppose that the minimal inter-pulse interval at which the listener could detect the test pulse (Tmin) can define the temporal resolution of pulse loudness. Another parameter, the shift of pulse detection thresholds (dIso) may characterize the change of test pulse loudness in pulse train.

The temporal resolution of pulse loudness equal to 25 ms for the typical listeners with normal hearing (Lt) corresponded to the amplitude modulation frequency threshold for noise, which obtained using temporal modulation transfer function (6). The auditory system is considered to analyze amplitude changes of sound signals if the modulation frequency is less than the modulation threshold. The listeners with aging-related hearing loss have lower sensitivity to amplitude modulation (7-10). In our study, the listener with aging-related hearing loss (L2) has the high Tmin.

The temporal integration of tones in our listeners with normal hearing (Lt and L5) and with small hearing losses (L2) corresponded to the well-known data (11-14).

The temporal integration of pulse trains obtained for "typical" listeners (Lt) with normal hearing, revealed known properties (15), but the individual data of the two other listeners (I5 and I2) with different auditory sensitivity did not. For them, the indicator of I_ps at T of 20 ms was 6-8 dB. The auditory system of these listeners could unite the responses on the pulses in train, therefore the detection thresholds of the train was less than the detection threshold of a single pulse.

The listeners of L5 and L2 had the different values of Tmin for different positions of test pulses in train. In the case of different positions of the test pulse, the shifts of dIso obtained for inter-pulse interval equal to Tmin could exceed 20 dB, but connections between the indicators of Tmin and the shifts of dIso were different. These pointed out to the differences in the processes of the loudness formation for first and last pulses.

For the first pulse in the train, the following rule was true: the higher the value of Tmin the greater the shift dIso. Apparently, the relatively short-term integration process (with time constant less than

110 ms) participated in the formation of the first pulse loudness. This process was responsible for the partial masking of the pulse loudness (4, 5).

For the last pulse in the train, another rule was valid: the smaller the value of Tmin, the greater the shift of dIso. We supposed, in addition to the short-term integration process, a long-term integration process of the post-spike changes in the auditory neuron excitability participated in the formation of the last pulse loudness. This long-term process could be responsible for both the forward masking (5) and changes in loudness of last pulse. The contribution of long-term process to the last pulse detection could be considerably greater than the contribution of the short-term process.

Thus, the detectability of the threshold loudness of the first and last pulses in train depended on the inter-pulse interval. Due to this it is possible to suppose that the phoneme loudness may depend on speech rate. Decrease phoneme loudness may be a reason for a decrease in speech intelligibility.

ACKNOWLEDGEMENTS

Supported by the Russian Foundation for Basic Research (RFBR, grants No 14-04-00155 and 17-04-00096).

REFERENCES

1. Fitzgibbons PJ, Gordon-Salant S. Age-related differences in discrimination of temporal intervals in accented tone sequences. Hearing Research. 2010. 264. p. 41–47.
2. Fitzgibbons PJ, Gordon-Salant S. Age effects in discrimination of repeating sequence intervals. J Acoust Soc Am. 2011. 129. p. 1490-1500.
3. Miller J, Aibel IL, Green K. On the nature of rate-dependent processing during phonetic perception. Perception and Psychophysics. 1984. 35. p. 5-15.
4. Fastl H. Temporal partial masking of pure tones by broad-band noise: Experimental result and models. Acta Acustica united with Acustica. 1984. 54. p. 145-153.
5. Fastl H, Zwicker E. Psychoacoustics: Facts and Models. 2008. Springer.
6. Viemeister NF. Temporal modulation transfer functions based upon modulation thresholds. J Acoust Soc Am. 1979. 66. p.1364-138
7. Formby C. Modulation threshold functions for chronically impaired Meniere patients. Audiology. 1987. 26. p. 89-102.
8. Bacon SP, Viemeister NF. Temporal modulation transfer functions in normal-hearing and hearing-impaired listeners. Audiology. 1985. 24. p.117-134.
9. Bacon SP, Gleitman RM. Modulation detection in subjects with relatively flat hearing losses. Journal of Speech and Hearing Research, 1992. 35. p., 642-653.
10. Moore BCJ, Shailer MJ, Schooneveldt GP. Temporal modulation transfer functions for band-limited noise in subjects with cochlear hearing loss. British Journal of Audiology. 1992. 26. p. 229-237
11. Green DM, Birdsall TG, Tanner WP. Signal detection as a function of signal intensity and duration, J Acoust Soc Am. 1957. 29. p. 523–531.
12. Zwislocki JJ. Theory of temporal auditory summation. J. Acoust. Soc. Am. 1960, 32, P.1046-1060;
13. Elliott LL. Temporal and masking phenomena in persons with sensorineural hearing loss. Audiology. 1975. 14. p. 336–353.
14. Florentine M, Fastl H, Buus S. Temporal integration in normal hearing, cochlear impairment, and impairment simulated by masking. J Acoust Soc Am. 1988. 84. p. 195-203.
15. Garner WR, Miller GA. The masked thresholds of pure tones as a function of duration. J Exp Psychol. 1947. 37. p. 293-303.