DOI 10.1007/s11055-017-0510-3
Neuroscience and Behavioral Physiology, Vol. 47, No. 9, November, 2017

Effects of Russian-Language Word Frequency on Mismatch Negativity in Auditory Event-Related Potentials

A. A. Aleksandrov, K. S. Memetova,
L. N. Stankevich, and K. O. Uplisova

Translated from Rossiiskii Fiziologicheskii Zhurnal imeni I. M. Sechenova, Vol. 102, No. 6, pp. 717–728, June, 2016. Original article submitted February 20, 2016. Revised version received April 6, 2016.

We report here our studied on the influence of the lexical frequency of Russian words on the latency and amplitude of mismatch negativity (MMN) in auditory event-related potentials (ERP). ERP were recorded in a multideviant oddball paradigm by presenting different-frequency Russian words and pseudowords. These experiments showed that the pattern of intrinsic MMN differed significantly between words with different lexical frequencies ($p = 0.001$) – the higher the frequency, the greater the amplitude and the shorter the latent period of the intrinsic MMN of the words. It is suggested that the greater amplitude and shorter latency of MMN for high-frequency words as compared with the pattern of MMN for low-frequency words is due to activation of memory traces for these words, these being stored in the cerebral cortex as distributed neuron populations. The suggestion that there is superfast access to lexical information during speech perception is confirmed, with access being possible 100–200 msec after presentation of a word. The ratio of MMN amplitudes for different pseudowords was somewhat reminiscent of data on MMN for words (analogs of high-frequency words produced higher-amplitude responses, while analogs of low-frequency words produced weaker responses, with no significant difference between low- and intermediate-frequency analogs), though MMN amplitudes for pseudowords were significantly greater and latent periods were significantly longer. Increases in the amplitude and latency of MMN to pseudowords as compared with MMN to words is associated with later and uncertain recognition of rarely encountered low-frequency words and completely unfamiliar stimuli, which are later classified as signals of a different category.

Keywords: event-related potentials, word frequency, speech perception, mismatch negativity.

Speech perception processes are undoubtedly among the most complex mechanisms of functioning in the human brain. The nature of linguistic representation in the brain remains an enigma despite an enormous number of studies on this question. The development of methods for recording the processes of brain operation has allowed us to approach a solution to the questions of the neurophysiology of speech and language. EEG methods provide for recording of event-related potentials (ERP) in the human brain with high time resolution during speech perception.

A particular method for studying the neural mechanisms underlying speech functions is provided by the early component of event-related potentials – mismatch negativity [20, 24]. The first data using auditory MMN, confirming the existence of memory traces for language stimuli, were obtained by Näätänen et al. [19] using materials in Finnish and Estonian. These authors showed that two mechanisms operate in parallel during speech perception: 1) detection of acoustic changes, which involves both hemispheres of the brain; 2) detection of phonematic changes, involving mainly the left hemisphere. Dehaene-Lambertz [7] also observed linguistic memory traces, by recording auditory ERP in the MMN paradigm to detect changes in MMN amplitude depending on changes in syllables. These authors provided the first recordings [3] of the effects of word frequency on the pattern of auditory MMN using a reversive oddball paradigm. These studies showed that high-frequency words gen-

---

Faculty of Biology, St. Petersburg State University,
St. Petersburg, Russia; e-mail: k.memetova@spbu.ru.

TABLE 1. Word Frequencies, Data from the New Frequency Dictionary of the Russian Lexicon and the Frequency Dictionary of the Living Russian Language

Word	Frequency of lemma, MWF (per million word forms)	Word frequency in living Russian speech, MWF
Chas	643.82	716.9
Chai	145.62	307.4
Chan	5.2	–

erated auditory MMN with higher amplitude and shorter latency than low-frequency words. The authors put forward the hypothesis that memory traces for high-frequency words, stored in the cerebral cortex as interconnected distributed neuron populations, are activated. Word frequency is a measure of the frequency with which a word is used in native speech. The results provided evidence supporting the view that frequently used or high-frequency words are processed differently from rarely encountered words, low-frequency words, or pseudowords. It is suggested that speech perception activates distributed neural networks associated with word representations, the frequency of use of lexical units affecting these neural connections. In the case of frequently used (high-frequency) words, strengthening of the neural representation leads to the generation of intrinsic MMN with greater amplitude and/or shorter latency.

In this study, Russian-language material was used to test whether word frequency affects "intrinsic" MMN. It has repeatedly been suggested that there is an increase in MMN amplitude compared with the standard in responses to rare deviant stimuli because of differences in the "refractoriness" (stimulus-specific adaptation) of neural populations tuned to the acoustic properties of the standard and deviant. Comparison with responses to deviant stimuli, or responses to physically identical stimuli presented among a multitude of equally probable sounds with the same probability as deviants in the oddball paradigm, or responses to physically identical stimuli presented with the same probability as the standard in the oddball paradigm has been proposed as an approach to answering this question [18, 21, 28]. Use of these methods has confirmed the existence of "intrinsic" MMN [1]. It should be noted that the optimum ratio of standard and deviant stimulus presentation probabilities decreasing the contribution of the refractoriness of neurons to generating mismatch negativity, according to published results [1], is a ratio of 85% standard to 15% deviant stimuli. Our study compared intrinsic MMN waves produced in responses to high-, intermediate-, and low-frequency words presented in a passive multistimulus oddball paradigm. The paradigm used here was developed by Näätänen et al. [20]. It has the feature of allowing presentation of one standard and several deviant stimuli, decreasing the duration of the experiment and increasing the set of stimuli presented. A special role was assigned to selection of words with different frequencies, as it was important to define their frequencies in conversational speech [2]. For the word frequency effect to influence MMN parameters in addition to detection of minimal word changes [21], three different-frequency words with maximal similarity to each other in terms of their physical properties (duration, spectral characteristics, etc.) were selected. In addition, an additional version of the test was run, which included pseudowords constructed in compliance with these rules and analogous to real Russian words to study the possible effects of the acoustic properties of words on MMN.

Methods. Two experimental variants were performed: A) with presentation of Russian-language words with different frequencies and B) presentation of pseudowords constructed according to the rules of the Russian language and constituting analogs of the words used in variant A. Each study variant involved 10 healthy subjects with good hearing, who stated that they were right-handed, aged 23–28 years, and who were native Russian-speakers. The study was approved by the St. Petersburg State University Ethics Committee.

In variant A, stimuli were words with different frequencies in Russian speech. Three words were selected, corresponding to the CVC formula (consonant–vowel–consonant), whose frequencies were determined using the New Frequency Dictionary of the Russian Lexicon and the Frequency Dictionary of Living Russian Speech [2]; values are given in Table 1.

Each subject was presented with three sets of stimuli in pseudorandom order: 1) the low-frequency word ***chan***[1] as a standard stimulus and ***chas*** and ***chai*** (high- and intermediate-frequency words) as deviant stimuli; 2) the intermediate-frequency word ***chai*** as standard and ***chas*** and ***chan*** (high and low frequency, respectively) as deviant stimuli; 3) the high-frequency word ***chas*** as standard stimulus and ***chai*** and ***chan*** (intermediate and low frequency, respectively) as deviants.

In variant B, stimuli were pseudowords: ***shas***, an analog of the high-frequency word ***chas***, ***shai***, an analog of the intermediate-frequency word ***chai***, and ***shan***, an analog of the low-frequency word ***chan***. Each subject was presented

---

[1] **Translator's note.** The Russian words *chan*, *chas*, and *chai* mean *vat*, *hour*, and *tea*, respectively. In all three words the element *cha* is pronounced as *cha* in the English word *chat*. The same applies to the pseudowords, *shan*, *shas*, and *shai*, none of which has any meaning in Russian.

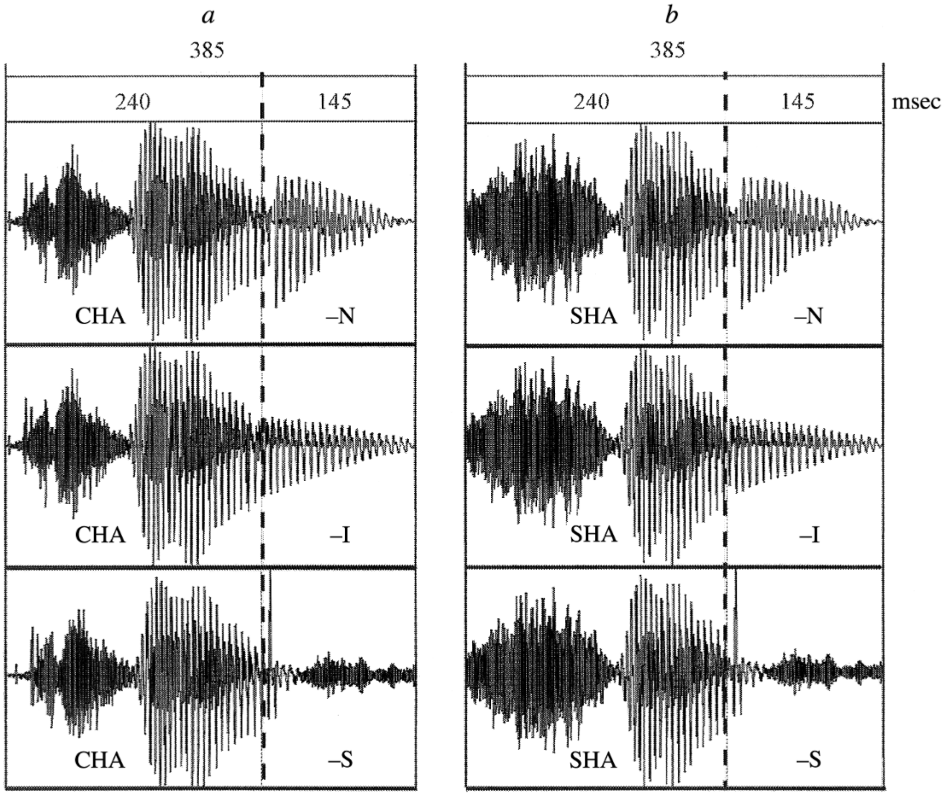

Fig. 1. Oscillograms of stimuli: *a*) words; *b*) pseudowords. Horizontal axes show stimulus duration, msec. Total stimulus duration was 385 msec; stimulus beginnings (the first two phonemes) lasted 240 msec; dashed lines show the divergence points of the stimuli; the last phoneme, responsible for the differences between the stimuli, lasted 145 msec.

with three sets of stimuli in pseudorandom order: 1) ***shan*** as the standard stimulus and ***shas*** and ***shai*** as deviant stimuli; 2) ***shai*** as the standard and ***shas*** and ***shan*** as deviant stimuli; 3) ***shas*** as the standard stimulus and ***shai*** and ***shan*** as deviant stimuli.

The study used a passive multistimulus oddball paradigm in which deviant stimuli (D_x, D_y) were distributed among standard stimuli (S). The probability that a deviant stimulus would appear among standard stimuli was determined by the ratio 85% S to 15% D, which has been indicated in [1, 11] to be optimum in terms of decreasing the contribution of neuron refractoriness to the generation of mismatch negativity.

Stimuli were synthesized using Acapela Group Virtual Speaker (female voice) Duration was 385 msec. Each study variant involved presentation of stimuli differing from each other only in terms of the last phoneme. The point at which the last phoneme was substituted was termed the divergence point. The interval from the beginning of stimulus presentation to the divergence point was 240 msec (Fig. 1). Thus, the stimuli were absolutely identical from the start of presentation to the divergence point. The poststimulus interval was 500 msec with randomization to 50 msec. Totals of 1334 stimuli in pseudorandom order were presented in each condition.

The physical properties of stimuli (amplitude, duration, intensity, spectral properties) were as similar as possible. Thus, the acoustic contrast between the standard and the deviant was identical in all three combinations.

The total duration of the experiments was 90 min. During this time, three sets of stimuli from variant A and three sets of stimuli from variant B were presented to the subject each of about 25 min duration. The subject 's task was to remain as relaxed as possible and watch a silent video on the monitor screen positioned opposite, without any requirement to pay attention to the stimuli presented.

EEG recording. During the study, subjects were placed in an acoustically insulated room and watched a soundless video on a monitor screen. Words were presented using the program Presentation binaurally via headphones at a comfortable sound level (50 dB). EEG recordings were made using silver chloride electrodes positioned on the surface of the head in leads F3, Fz, F4, C3, Cz, and C4 on the international 10–20 scheme [12]. The reference electrode was located on the tip of the nose and the ground electrode was placed on the forehead. Electrical artifacts evoked by eye movements were tracked by recording the electrooculogram. Electrode resistance was no greater than 5 kΩ. Signals were digitized at a sampling frequency of 250 Hz and were

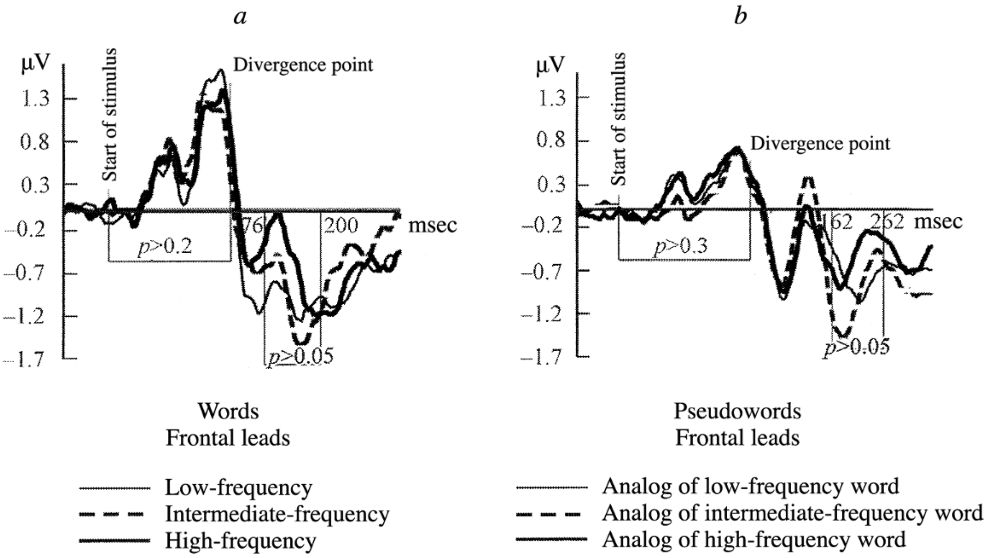

Fig. 2. Event-related potentials to stimuli presented as standards: *a*) words; *b*) pseudowords. Overall results for frontal leads are shown (number of subjects = 10). Horizontal axes show ERP latent periods, msec; vertical aces show ERP response amplitudes, μV. Traces of ERP waves correspond to stimuli as follows: thin line – low-frequency word and its analog; dotted line – intermediate-frequency word and its analog; thick line – high-frequency word and its analog.

filtered in the band 1–30 Hz. The isoline was corrected using the 100 msec prestimulus interval. Epochs in which the EEG and EOG signals were greater than 100 μV were regarded as artifacts and were eliminated from the analysis. ERP data in which the number of artifacts for deviant or standard stimuli was greater than 15% were also excluded from subsequent group analysis.

EEG recordings were made using a Mitsar 24-channel digital encephalograph (bandpass 0.05–70 Hz) and WinEEG software for recording and processing electroencephalograms (V. A. Ponomarev, Bekhtereva Institute of the Human Brain, Russian Academy of Sciences). This experimental design provides for obtaining and analyzing intrinsic MMN, as one and the same stimulus is presented at the probability of the standard and also at the probability of the deviant. Intrinsic MMN was calculated as the difference between the responses to a given stimulus presented as deviant and as standard. Peak latency was calculated individually for each subject. The MMN peak was identified as the high-amplitude negative wave with latency 100–200 msec. For statistical analysis, an interval of 76–200 msec was selected for MMN in variant A and an interval of 162–262 msec for variant B. Results were assessed statistically by analysis of variance (ANOVA) in SPSS using the Bonferroni correction. A three-factor model with the factors "Stimulus type" (number of levels 3: low-, intermediate-, and high-frequency stimuli), "Condition" (number of levels 2: standard and deviant), and "Lead" (number of levels 6: F3, Fz, F4, C3, Cz, and C4) was used.

Results. *Variant A with different-frequency Russian-language words.* Comparison of ERP to stimuli presented as standards revealed no significant differences from the beginning of presentation to the divergence point ($p > 0.2$ for the frontal leads (Fig. 2, *a*). After the divergence point, there were also no differences in evoked potentials, including in the interval 76–200 msec ($p > 0.05$), which was the period with the greatest spread of values.

Analysis of variance for repeat measures of ERP to deviant stimuli revealed significant influences of the factors "Stimulus type" ($F_{(1,762)} = 7.682$, $p = 0.003$), "Condition" ($F_{(1,872)} = 5.871$, $p = 0.004$), and the interaction of the "Stimulus type" × "Condition" factors ($F_{(1,754)} = 6.185$, $p < 0.004$) on the amplitude and latency of responses in the interval 76–200 msec. Pairwise comparison showed that the amplitude and latency of the wave for the deviant high-frequency word ***chas*** was significantly different from the amplitude and latency of the wave for the deviant intermediate-frequency word ***chai*** ($p = 0.001$) and the amplitude of the wave for the deviant low-frequency word ***chan*** ($p = 0.006$); significant differences were also seen between deviant responses for the pair of intermediate- and low-frequency words ***chai*** and ***chan*** ($p = 0.022$).

As a result, determination of intrinsic MMN for words (calculated as the differences between responses to the same stimulus presented as deviant and as standard) revealed significant effects on the amplitude and latency of intrinsic MMN in the interval 76–200 msec for the factors "Stimulus type" ($F_{(1,771)} = 104.753$, $p = 0.001$), "Condition" ($F_{(1,824)} = 36.136$, $p = 0.001$), and the interaction of the "Stimulus type" and "Condition" factors ($F_{(1,782)} = 54.881$, $p = 0.001$) (Fig. 3). Pairwise comparisons showed that the amplitude and latency of the intrinsic MMN of the high-frequency word ***chas*** were significantly different from the amplitude and latency

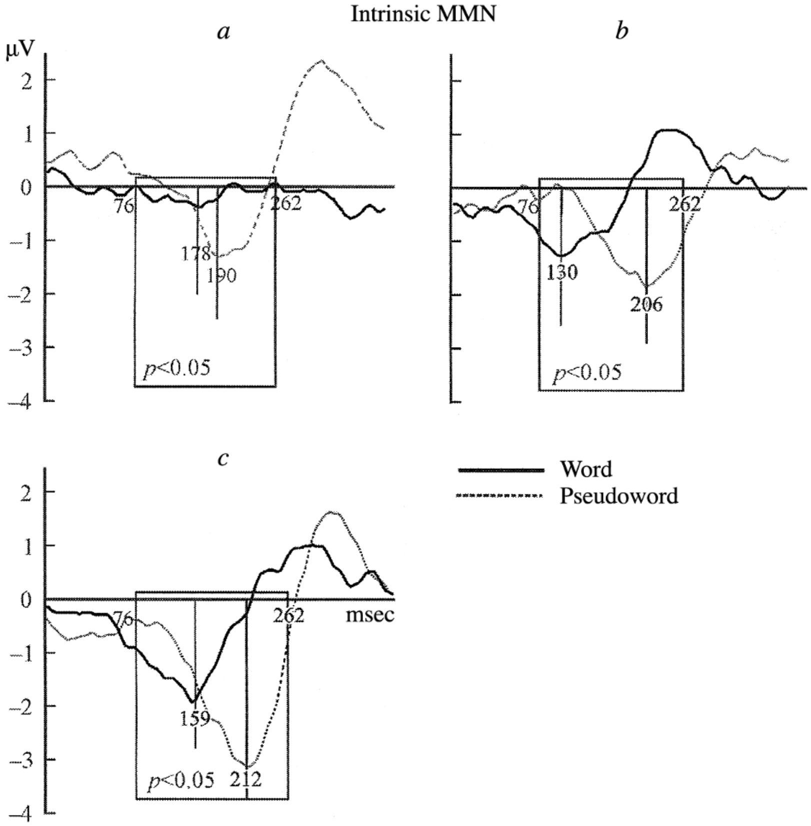

Fig. 3. Intrinsic MMN for words and pseudowords: *a*) low-frequency word and its analog; *b*) intermediate-frequency word and its analog; *c*) high-frequency word and its analog. Overall results for the frontal leads are shown (number of subjects = 10). The horizontal axes show latent periods of ERP responses, msec; the vertical axes show the amplitude of ERP responses, μV. The maximum intrinsic MMN amplitude peak is identified by the marker, with indication of latency, msec. Continuous lines show MMN waves for the word; dotted lines show MMN waves for pseudowords.

of the intrinsic MMN of the intermediate-frequency word ***chai*** ($p = 0.001$) and also from the amplitude and latency of the MMN of the low-frequency word ***chan*** ($p = 0.001$). In addition, there were also significant differences between the intrinsic MMN of pairs of intermediate- and low-frequency words ***chai*** and ***chan*** ($p = 0.001$).

Variant B with pseudowords. Comparison of ERP to stimuli presented as pseudowords revealed no significant differences (before the divergence point) ($F_{(2)} = 0.438, p > 0.3$) or after the divergence point in the interval 162–262 msec ($F_{(2)} = 4.510, p > 0.05$) (Fig. 2, *b*).

Analysis of variance for repeat measures of ERP to deviant stimuli demonstrated significant influences of the “Condition” factor ($F_{(1,127)} = 40.962, p = 0.001$) and the interaction of the “Stimulus type” and “Condition” factors ($F_{(1,386)} = 28.819, p = 0.001$) on the amplitude and latency of responses in the interval 162–262 msec. Pairwise comparison showed that the amplitude and latency of the wave of the deviant high-frequency word analog ***shas*** were significantly different from the amplitude and latency of the wave for the deviant intermediate-frequency word analog ***shai*** ($p = 0.001$), and from the amplitude and latency of the wave for the deviant low-frequency word analog ***shan*** ($p = 0.002$). In addition, significant differences were also seen between deviant responses for the pair consisting of the intermediate- and low-frequency pseudowords ***shai*** and ***shan*** ($p = 0.001$).

As a result, determination of intrinsic MMN for pseudowords revealed significant influences for the “Stimulus type” factor ($F_{(1,243)} = 26.776, p < 0.037$) and the interaction of the “Stimulus type” and “Condition” factors ($F_{(1,476)} = 7.185, p < 0.003$) on response amplitude and latency in the interval 162–262 msec (Fig. 3). Pairwise comparisons showed that the amplitude and latency of the intrinsic MMN of the high-frequency word analog ***shas*** were significantly different from the amplitude and latency of the intrinsic MMN of the low-frequency word analog ***shan*** ($p = 0.014$). There were no significant differences between the amplitude and latency of the intrinsic MMN of the intermediate-frequency word analog ***shai*** and the other two analogs (of high- and low-frequency words), ***shas*** and ***shan*** ($p = 0.859$ and $p = 0.106$, respectively).

Comparison of the intrinsic MMN of words and the intrinsic MMN of pseudowords using analysis of variance for repeat measures revealed significant effects of

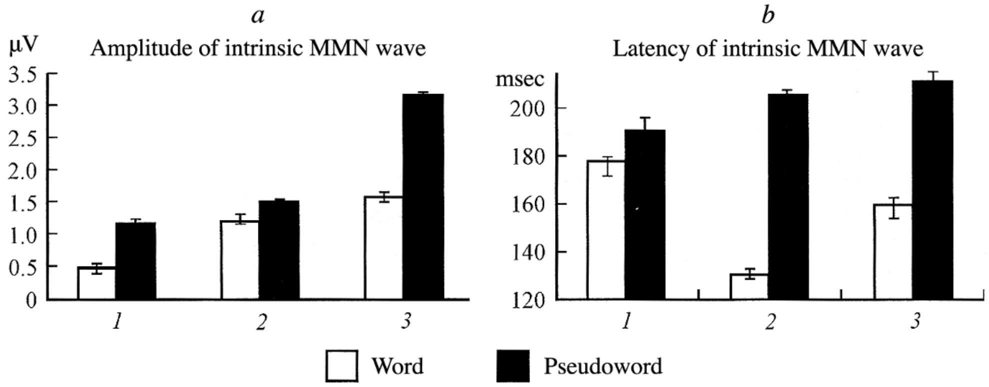

Fig. 4. Histograms comparing intrinsic MMN waves for stimuli: *a*) intrinsic MMN amplitude; *b*) intrinsic MMN latency. Overall results are shown for the frontal leads (number of subjects = 10). The horizontal axes show stimuli: *1*) low-frequency word and its analog; *2*) intermediate-frequency word and its analog; *3*) high-frequency word and its analog (white columns show words; dark columns show pseudowords); the vertical axis in *a* shows MMN response amplitude, μV, and the vertical axis in *b* shows MMN response latency, msec.

the "Condition" factor ($F_{(1,000)}$ = 12.911, p = 0.002), the "Stimulus type" factor ($F_{(1,438)}$ = 369.360, p = 0.001), and the interaction of the "Stimulus type" × "Condition" factors ($F_{(1,114)}$ = 3.607, p = 0.05) on MMN amplitude and latency. Pairwise comparisons revealed differences between the amplitudes and latent periods of the intrinsic MMN of words and the amplitudes and latent periods of the intrinsic MMN of pseudowords (p = 0.002). The amplitude of the intrinsic MMN to pseudowords was significantly greater than the amplitude of intrinsic MMN to words, for example a more than two-fold difference in the case of the low-frequency word and its analog (Fig. 4, *a*). The latency of the intrinsic MMN to the pseudoword was also significantly greater than the latency of the MMN peak for the word, the differences between mean values reaching 53 and 76 msec for the MMN of high-frequency and intermediate-frequency words, respectively (Fig. 4, *b*).

The absence of any significant interaction between the "Stimulus type" × "Condition" × "Lead" factors in the two blocks is evidence that the potentials recorded to the deviant and standard stimuli had identical topographic distributions across the brain surface.

Discussion. Analysis of MMN for words showed a significant increase in MMN amplitude in the period 76–200 msec during processing of the high-frequency word as compared with the intermediate- and low-frequency words, as well as during processing of the intermediate-frequency word as compared with the low-frequency word. Changes in the amplitude of the MMN wave cannot be explained solely by their acoustic differences, as these were minimal, but probably represent detection of the effects of word frequency. A previous study [3] showed that MMN amplitude on processing of a high-frequency word was markedly greater than MMN amplitude for a low-frequency word and that MMN latency for a low-frequency word was significantly greater than the latency for a high-frequency word. The authors used two different-frequency words as stimuli, though pseudowords were not used. These data are consistent with the results obtained in our study. We showed that the pattern of intrinsic MMN differed significantly between words of different frequencies (p = 0.001): the greater the frequency, the greater the amplitude and the shorter the latent period of intrinsic MMN. It is important to note that in the case of the low-frequency word, determination of the latency of the intrinsic MMN can be quite arbitrary, as the amplitude of the deviation of the MMN wave was low and local low-amplitude peaks barely differed from each other.

Thus, testing of the hypothesis that word frequency affects MMN parameters in auditory event-related potentials using a different, multideviant, paradigm with calculation of intrinsic MMN for entirely different pairs of words with different frequencies confirms the initial hypothesis [3].

In the present study, the possible influence of minimal acoustic differences on intrinsic MMN was evaluated using pseudowords differing from words only in the first phoneme and were subsequently identical in terms of acoustic properties to words. Comparison of ERP for pseudowords showed that the pattern of the intrinsic MMN differed significantly only in one case: between the pair ***shan-shas*** (p < 0.015), while there was no significant difference for the pairs ***shas-shai*** and ***shan-shai***.

Another important characteristic of MMN for pseudowords consisted of significant differences from MMN for words in terms of both the amplitude and the latency of the MMN peak. It should be noted that the latent periods of intrinsic MMN for pseudowords were in a completely different time range. They were significantly longer than those of intrinsic MMN for words, and were in the range 162–262 msec, as compared with 76–200 msec. The amplitudes of intrinsic MMN waves to the pseudoword were significantly greater than the amplitude of the intrinsic MMN wave to the word. As regards the ratio of MMN amplitudes for different pseudowords, these were in some sense reminiscent of data on MMN for words (the analog of the high-frequency word produced the response with the greatest amplitude, while the analog of the low-frequency word produced the weakest response), though there were no significant differences between the pair ***shan-shas*** or between the pairs ***shas-shai*** and ***shan-shai***.

Secondly, the amplitudes of intrinsic MMN waves to pseudowords were significantly greater than the amplitudes of intrinsic MMN waves to words. All this suggests that our data on the characteristics of MMN to words and pseudowords cannot be explained only in terms of the acoustic features of the stimuli. To explain the significant increases in MMN amplitude and latency when pseudowords were used, we should remember that the appearance of a deviant with characteristics sharply different from those of standards has in some studies produced sharp increases in neurons (responses to so-called novel stimuli) [17, 18]. It may be that presentation of verbal stimuli consisting of pseudowords is linked with slow processing and failure to complete semantic identification. This leads on the one hand to an increase in the duration of the process and, on the other, to mismatch negativity being generated not to the contrast of high-frequency/low-frequency words but to the word/pseudoword contrast, resulting in a sharp increase in MMN latency to pseudowords and a significant increase in MMN amplitude, which cannot be explained solely in terms of the acoustic characteristics of the stimuli. It can be suggested that we observed a tendency to a later and imprecise recognition of rarely encountered low-frequency words or completely unfamiliar stimuli, which take longer to classify. In this regard, it has to be recognized that the pseudowords used as a control procedure in this paradigm cannot be an optimum solution, considering the results obtained on the properties of pseudoword perception and the fact that that it could potentially be affected by the similarity of a pseudoword to one word or another (with different frequencies).

Overall, these data support the hypothesis that word frequency influences MMN amplitude and latency. In fact, superfast neural access to lexical information has been demonstrated, possibly by 100–200 msec [24]. Such an early cerebral response to lexical stimuli can be explained by the distributed neural network present in the brain – the "word network" [23], which forms during the language acquisition process. The greater MMN amplitude to high-frequency words as compared with low-frequency words may reflect activation of long-term memory traces. It would seem that we are dealing with a widely distributed network of neuronal ensembles, involved in processing speech and language. Instantaneous activation of such a network leads to a rapid, almost simultaneous activation of all its connections, which we see as MMN. The magnitude of this type of cerebral response must depend on the strength of the internal connections formed by memory [13, 24, 28]. The distributed neural "word network" is thus due to at least two speech processes: perception and motor function. Other mechanisms required for processing incoming information (presentation modality, semantics of the word being processed, etc.) are also added in. Such "word networks" are formed by short and long neural connections, whose complexity depends on the quantity and frequency of activation. When neurons are activated simultaneously, the synaptic connections between them are strengthened [3, 33]. Thus, frequently used words lead to strengthening of neuron connections active during processing, as there is constant simultaneous activation of neurons in those parts of the cortex involved. These concepts provide an explanation for the differences seen in intrinsic MMN for different-frequency words.

This study was supported by the Russian Humanities Scientific Foundation (Project No. 15-06-10806).

REFERENCES

1. M. D. Evstigneeva and A. A. Aleksandrov, "Mismatch negativity: contribution of differences in the refractoriness of stimulus-specific neuronal populations," *Ros. Fiziol. Zh.*, **94**, No. 10, 1147–1157 (2008).
2. O. N. Lyashevskaya and S. A. Sharov, *Frequency Dictionary of the Russian Language (data from the Russian National Corpus)*, Azbukovnik, Moscow (2009).
3. A. A. Alexandrov, D. O. Boricheva, F. Pulvermuller, and Y. Shtyrov, "Strength of word-specific neural memory traces assessed electrophysiologically," *PLoS One*, **6**, (2011); doi: 10.1371/journal.pone.0022999.
4. R. Assadollahi and F. Pulvermuller, "Neuromagnetic evidence for early access to cognitive representations," *Neuroreport*, **12**, No. 2, 207–213 (2001).
5. S. Boudelaa, F. Pulvermuller, O. Hauk, et al., "Arabic morphology in the neural language system," *J. Cogn. Neurosci.*, **22**, 998–1010 (2010).
6. C. M. Brown, P. Hagoort, and M. ter Keurs, "Electrophysiological signatures of visual lexical processing: open- and closed-class words," *J. Cogn. Neurosci.*, **11**, No. 3, 261–281 (1999).
7. G. Dehaene-Lambertz, "Electrophysiological correlates of categorical phoneme perception in adults," *NeuroReport*, **8**, 919–924 (1997).
8. G. Dehaene-Lambertz and S. A. Baillet, "A phonological representation in the infant brain," *NeuroReport*, **9**, 1885–1888 (1998).
9. D. Embick, M. Hackl, J. Schaeffer, et al., "A magnetoencephalographic component whose latency reflects lexical frequency," *Cogn. Brain Res.*, **10**, 345–348 (2001).
10. T. Endrass, B. Mohr, and F. Pulvermuller, "Enhanced mismatch negativity brain response after binaural word presentation," *Eur. J. Neurosci.*, **19**, 1653–1660 (2004).
11. M. D. Evstigneeva and A. A. Aleksandrov, "Mismatch negativity: The contribution of differences in the refractoriness of stimulus-specific neuron populations," *Neurosci. Behav. Physiol.*, **39**, No. 9, 833–840 (2009).
12. H. H. Jasper, "The ten-twenty electrode system of the International Federation of Societies for Electroencephalography: Appendix to report of the committee on methods of clinical examination in electroencephalography," *Electroencephalogr. Clin. Neurophysiol.*, **10**, 371–375 (1958).
13. Garagnani M, Y. Shtyrov, and F. Pulvermuller, "Effects of attention on what is known and what is not: MEG evidence for functionally discrete memory circuits," *Front. Hum. Neurosci.*, **3**, 10 (2009).
14. J. Hanna and F. Pulvermuller, "Neurophysiological evidence for whole form retrieval of complex derived words: a mismatch negativ-

ity study," *Front. Hum. Neurosci.*, **8**, 886 (2014); doi: 10.3389/fnhum (2014).00886.

15. O. Hauk and F. Pulvermuller, "Effects of word length and frequency on the human event-related potential," *Clin. Neurophysiol.*, **115**, 1090–1103 (2004).
16. O. Hauk, Y. Shtyrov, and F. Pulvermuller, "The sound of actions as reflected by mismatch negativity: rapid activation of cortical sensory-motor networks by sounds associated with finger and tongue movements," *Eur. J. Neurosci.*, **23**, 811–821 (2006).
17. J. Yue, R. Bastiaanse, and K. Alter, "Cortical plasticity induced by rapid Hebbian learning of novel tonal word-forms: Evidence from mismatch negativity," *Brain Lang.*, **139**, 10–22 (2014); doi: 10.1016/j.band1.09.007.
18. R. Näätänen and K. Alho, "Mismatch negativity – the measure for central sound representation accuracy," *Audiol. Neurootol.*, **2**, 341–353 (1997).
19. R. Näätänen, A. Lehtokoski, M. Lennes, et al., "Language-specific phoneme representations revealed by electric and magnetic brain responses," *Nature*, **385**, 432–434 (1997).
20. R. Näätänen, S. Pakarinen, T. Rinne, and R. Takegata, "The mismatch negativity (MMN): towards the optimal paradigm," *Clin. Neurophysiol.*, **115**, 140–144 (2004).
21. S. Pakarinen, R. Takegata, T. Rinne, et al., "Measurement of extensive auditory discrimination profiles using the mismatch negativity (MMN) of the auditory event-related potential (ERP)," *Clin. Neurophysiol.*, **118**, 177–185 (2007).
22. C. M. Pettigrew, B. E. Murdoch, C. W. Ponton, et al., "Automatic auditory processing of English words as indexed by the mismatch negativity, using a multiple deviant paradigm," *Ear Hear.*, **25**, 284–301 (2004).
23. F. Pulvermuller, T. Kujala, Y. Shtyrov, et al., "Memory traces for words as revealed by the mismatch negativity," *Neuroimage*, **14**, 607–616 (2001).
24. F. Pulvermuller and Y. Shtyrov, "Language outside the focus of attention: The mismatch negativity as a tool for studying higher cognitive processes," *Progr. Neurobiol.*, **79**, No. 1, 49–71 (2006).
25. M. Scharinger, P. J. Monahan, and W. J. Idsardi, "Linguistic category structure influences early auditory processing: Converging evidence from mismatch responses and cortical oscillations," *Neuroimage*, **128**, 293–301 (2016); doi:10.1016/j.neuroimage.01.003.
26. S. C. Sereno, C. C. Brewer, and P. J. O'Donnell, "Context effects in word recognition: evidence for early interactive processing," *Psychol. Sci.*, **14**, 328–333 (2003).
27. L. B. Shestopalova, E. A. Petropavlovskaia, S. P. Vaitulevich, and N. I. Nikitin, "Contextual effects on preattentive processing of sound motion as revealed by spatial MMN," *Int. J. Psychophysiol.*, **96**, No. 1, 49–56 (2015); doi:10.1016/j.ijpsycho.02.021.
28. Y. Shtyrov, T. Kujala, and F. Pulvermuller, "Interactions between language and attention systems: early automatic lexical processing?" *J. Cogn. Neurosci.*, **22**, 1465–1478 (2010).
29. Y. Shtyrov, L. Kimppa, F. Pulvermuller, and T. Kujala, "Event-related potentials reflecting the frequency of unattended spoken words: a neuronal index of connection strength in lexical memory circuits?" *Neuroimage*, **55**, 658–668 (2011).
30. Y. Shtyrov, G. Goryainova, S. Tugin, et al., "Automatic processing of unattended lexical information in visual oddball presentation: neurophysiological evidence," *Front. Hum. Neurosci.*, **7**, 421 (2013).
31. W. Sittiprapaporn, C. Chindaduangratn, M. Tervaniemi, and N. Khotchabhakdi, "Preattentive processing of lexical tone perception by the human brain as indexed by the mismatch negativity paradigm," *Ann. N. Y. Acad. Sci.*, **999**, 199–203 (2003).
32. A. Tavano, S. Grimm, J. Costa-Faidella, et al., "Spectrotemporal processing drives fast access to memory traces for spoken words," *Neuroimage*, **60**, 2300–2308 (2012).
33. T. Wennekers, M. Garagnani, and F. Pulvermuller, "Language models based on Hebbian cell assemblies," *J. Physiol. (Paris)*, **100**, 16–30 (2006).