Two-levels of feature binding in the auditory modality: reconciliation of opposing views

Boris V. Chernyshev[1], Dmitri V. Bryzgalov[1,2], Ivan E. Lazarev[1]

[1] Laboratory of Cognitive Psychophysiology,
[2] Centre for Cognition & Decision Making,
National Research University Higher School of Economics, Moscow, Russian Federation
(ilazarev@hse.ru)

Introduction

Feature binding is a hypothetical process that leads to holistic perception.

● Electrophysiological recordings of the mismatch negativity (MMN) show that detection of feature conjunctions is an early automatic process (Gomes et al., 1997, Sussman et al., 1998, Takegata et al., 2005, Winkler et al., 2005).

● Behavioral experiments evidence that feature binding is a non-automatic higher-level process (Treisman and Gelade, 1980, Thompson, 1994, Dyson and Quinlan, 2003).

Perception involving feature integration is now viewed as a non-linear network phenomenon involving multiple brain levels, thus casting doubt on feature binding being a single operation (Bullier, 2001). Moreover, since feature conjunctions may be specifically detected early in the processing stream (Bizley et al., 2009, Walker et al., 2011), representations of both single features and their conjunctions may be available at higher levels of processing. Thus, it has been supposed that the relative use of features themselves or their conjunctions may vary depending upon specific attentional task requirements, strategies used, etc. (Takegata et al., 2005).

The current study aimed at testing the hypothesis that processing of feature conjunctions involves both the early automatic level of preattentive sensory representations and the late integrative level leading to translation from sensory representations to a purposeful movement initiation. Particular focus in this study was on the analysis of mechanisms of accidental failures at both levels.

Methods

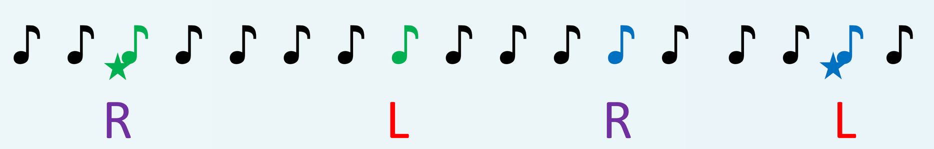
We used a 5-stimulus paradigm – a combination of a condensation task with an oddball task.

Four target auditory stimuli differed in complexity and pitch. Simple stimuli were sinusoidal tones; complex stimuli were made by adding a white noise to a sinusoidal tone. High and low tones had a frequency of 2000 and 1000 Hz correspondingly. Nontarget stimuli were sinusoidal 400 Hz tones.

Participants were instructed to make a two-alternative choice: to push one gamepad button in response to two target stimuli, and to push the other button to the other two target stimuli, while ignoring the nontarget stimuli.

	Simple Stimulus	★Complex stimulus
High	Left button	Right button
Low	Right button	Left button

All stimuli were presented in quasi-random sequences. The probability of the four target stimuli was equal; and they were otherwise equal as far as task requirements, attentional conditions, etc. may be concerned. The overall probability ratio of target to nontarget stimuli was 1:4.



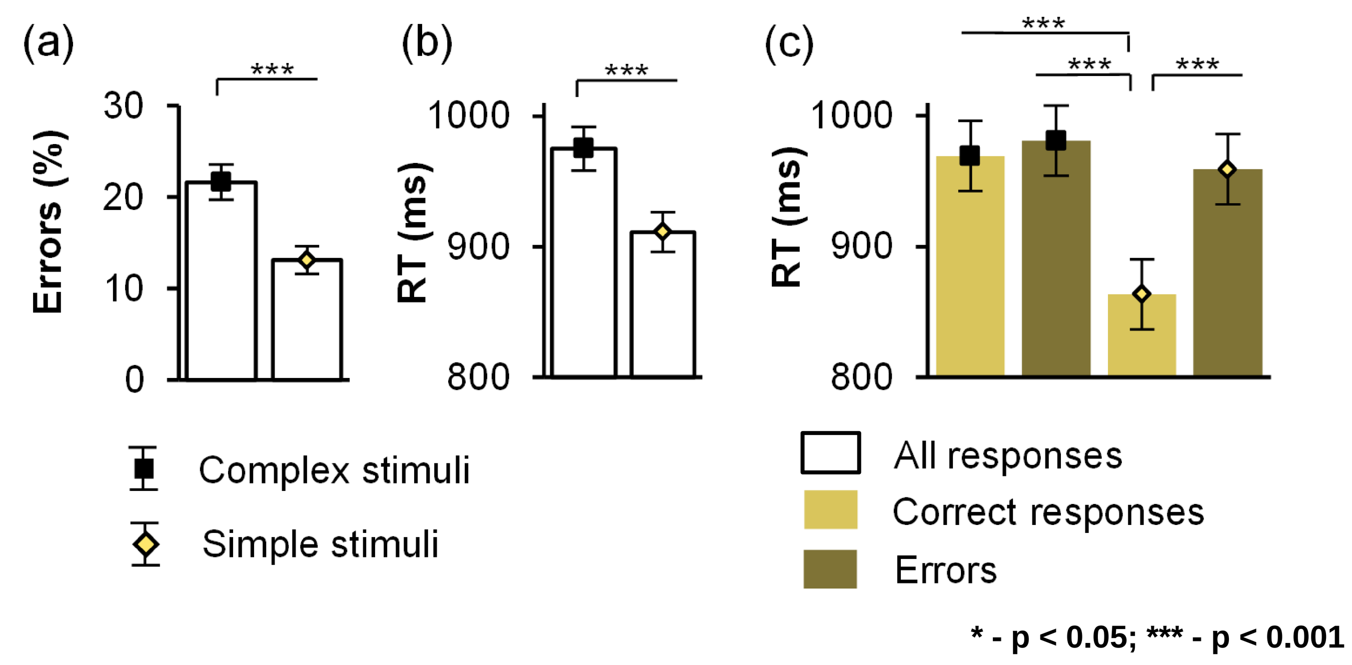
MMN amplitude was measured using differential ERPs by averaging over 120-220 ms time window. P300 was measured using target ERPs over 275-325 ms time window; these windows were chosen as time intervals during which the respective ERP components were most clearly pronounced. The following electrodes manifesting near-maximum amplitudes on scalp maps were included into the analysis: 9 electrodes for MMN (F3, Fz, F4, Fc3, Fcz, Fc4, C3, Cz, C4) and 8 electrodes for P300 (Cp3, Cpz, Cp4, P3, Pz, P4, T5, T6).

The behavioral data were compared using two tailed paired t-tests. ERP amplitudes were compared using repeated-measures ANOVAs with following factors: either Stimulus (two levels: complex and simple), or Response (two levels: correct and error), and Electrode (9 levels for MMN and 8 levels for P300). For each of the ERP comparisons, participants' data were included into analysis only on condition that each individual ERP was averaged across at least 5 trials or more.

The analyses reported below included 46 participants.

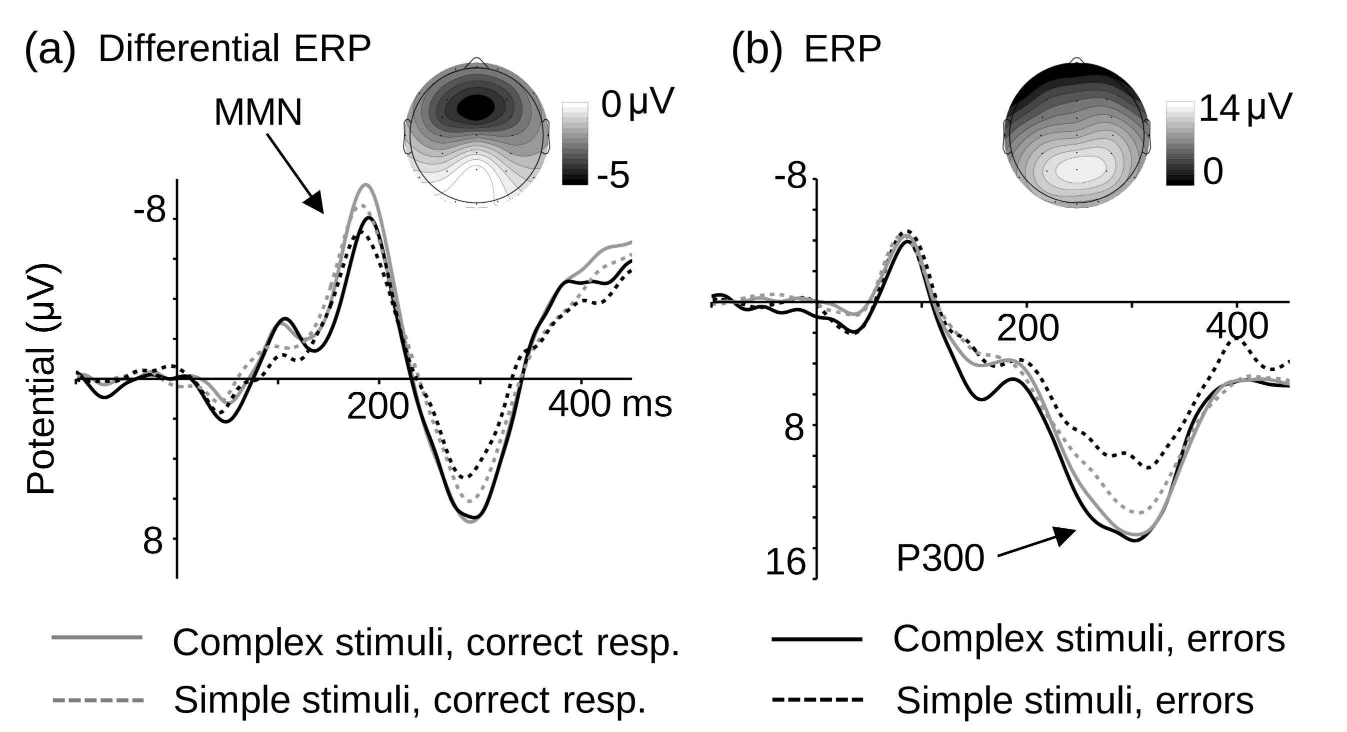
Results

Behavioral results:
percentage of errors (a) and response time (RT) (b, c).

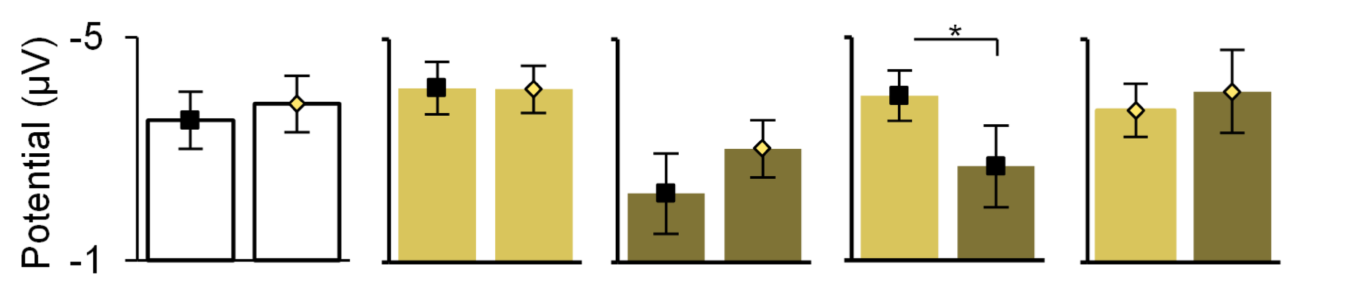


Grand average waveforms at Cz: differential ERP waveforms (deviant targets minus standards) (a) and ERP waveforms (b).

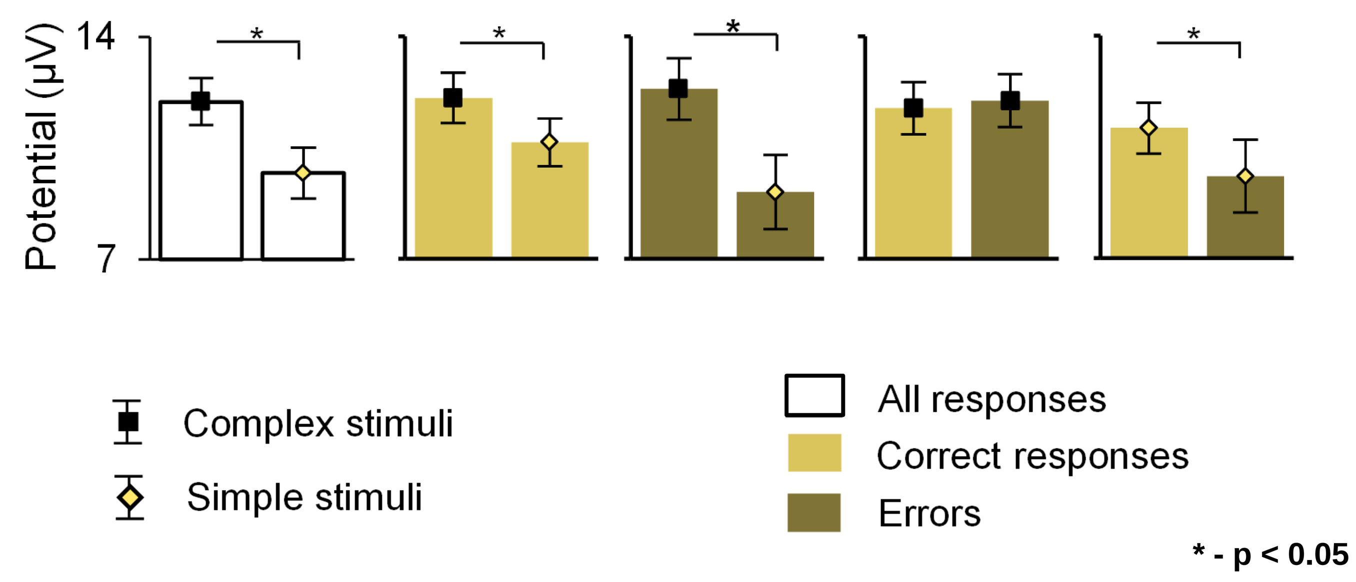
Insets: respective scalp maps averaged for both stimulus types and both response types over 120-220 ms and 277-325 ms correspondingly.



Mean MMN amplitude



Mean P300 amplitude



<u>**1.** MMN amplitude to complex stimuli did not differ from MMN amplitude to simple stimuli</u>. This finding stays in line with the MMN literature (Sussman et al., 1998, Oceak et al., 2008); thus the result of feature binding is available at the early preattentive level related to MMN generation.

<u>**2.** Responses to complex stimuli were slower and less accurate than responses to simple stimuli</u> – thus staying in line with behavioral reports (Thompson, 1994, Dyson and Quinlan, 2003) evidencing that binding occurs at higher level of sensory processing.

3. <u>On erroneous trials, MMN decreased in amplitude only for complex stimuli (but not for simple stimuli).</u> This observation hints that the neuronal population that encodes feature conjunctions is closely associated with (or coincides with) the neuronal population that generates MMN.

Thus, a critical stage of feature binding occurs during the early level related to MMN generation. Attentional distraction and arousal fluctuations were indeed shown to modulate MMN amplitude (Raz et al., 2001, Auksztulewicz and Friston, 2015), and behavioral errors were shown to reduce MMN amplitude (Elton et al., 2004).

Detection of pitch itself is hardwired at lower levels of auditory analysis, and accidental failures of automatic pitch discrimination seem very unlikely given the tone parameters used.

<u>4. P300 amplitude was greater for complex stimuli than for simple stimuli.</u> This stays in line with the observations that P300 increases in amplitude in response to deviant stimuli that differ from standards by a greater number of relevant features (Johnson, 1986, Lidji et al., 2010). This may indirectly evidence that complex stimuli recruit greater processing capacity (Kok, 2001). This is compatible with the hypothesis that higher processing levels receive sensory representations of conjoined features as well as of separate features.

Thus, the increased informational load created by complex stimuli may prolong RTs for the stimuli requiring conjunction – the effect that has been observed in the current study and was shown in earlier reports (Thompson, 1994, Dyson and Quinlan, 2003). In other words, the data evidence that, in addition to a low level, the analysis related to feature conjunction is also performed at higher levels.

<u>**5-6.** On erroneous trials, P300 amplitude was reduced for simple stimuli (but not for complex stimuli).</u>

<u>For simple stimuli, errors were associated with the delayed behavioral response, while no such delay was observed for complex stimuli.</u>

Reductions in P300 amplitude are known to be linked to deterioration in sensory processing and reduced performance (Polich, 2007, Smallwood et al., 2008).

Sensory pitch discrimination itself is not likely to cause errors to simple stimuli in the current task – thus the cause of rather rare errors in response to simple stimuli may be related to late higher-level neuronal processes, which serve to interpret the stimuli and to transform the sensory representations into preparation of appropriate motor responses (Verleger et al., 2005).

Conclusions

We found that feature binding in auditory modality involves at least two concurrent levels of processing.

This supports the notion that feature binding is a distributed multilevel process, involving both its detection at the level of the auditory cortex, and application of stimulus-to-response mapping at a higher integrative level.

References

Auksztulewicz R, Friston K (2015) Attentional Enhancement of Auditory Mismatch Responses: a DCM/MEG Study. Cereb Cortex 25:4273-4283.

Bizley JK, Walker KM, Silverman BW, King AJ, Schnupp JW (2009) Interdependent encoding of pitch, timbre, and spatial location in auditory cortex. The Journal of neuroscience : the official journal of the Society for Neuroscience 29:2064-2075.

Bullier J (2001) Integrated model of visual processing. Brain Res Brain Res Rev 36:96-107.

Dyson BJ, Quinlan PT (2003) Feature and conjunction processing in the auditory modality. Percept Psychophys 65:254-272.

Elton M, Spaan M, Ridderinkhof KR (2004) Why do we produce errors of commission? An ERP study of stimulus deviance detection and error monitoring in a choice go/no-go task. Eur J Neurosci 20:1960-1968.

Gomes H, Bernstein R, Ritter W, Vaughan HG, Jr., Miller J (1997) Storage of feature conjunctions in transient auditory memory. Psychophysiology 34:712-716.

Johnson R, Jr. (1986) A triarchic model of P300 amplitude. Psychophysiology 23:367-384.

Kok A (2001) On the utility of P3 amplitude as a measure of processing capacity. Psychophysiology 38:557-577.

Lidji P, Jolicœur P, Kolinsky R, Moreau P, Connolly JF, Peretz I (2010) Early integration of vowel and pitch processing: A mismatch negativity study. Neurophysiol Clin 121:533-541.

Oceak A, Winkler I, Sussman E (2008) Units of sound representation and temporal integration: a mismatch negativity study. Neurosci Lett 436:85-89.

Polich J (2007) Updating P300: An integrative theory of P3a and P3b. Neurophysiol Clin 118:2128-2148.

Raz A, Deouell LY, Bentin S (2001) Is pre-attentive processing compromised by prolonged wakefulness? Effects of total sleep deprivation on the mismatch negativity. Psychophysiology 38:787-795.

Smallwood J, Beach E, Schooler JW, Handy TC (2008) Going AWOL in the brain: mind wandering reduces cortical analysis of external events. Journal of cognitive neuroscience 20:458-469.

Sussman E, Gomes H, Nousak JM, Ritter W, Vaughan HG, Jr. (1998) Feature conjunctions and auditory sensory memory. Brain Res 793:95-102.

Takegata R, Brattico E, Tervaniemi M, Varyagina O, Näätänen R, Winkler I (2005) Preattentive representation of feature conjunctions for concurrent spatially distributed auditory objects. Cogn Brain Res 25:169-179.

Thompson WF (1994) Sensitivity to combinations of musical parameters: Pitch with duration, and pitch pattern with durational pattern. Percept Psychophys 56:363-374.

Treisman AM, Gelade G (1980) A feature-integration theory of attention. Cogn Psychol 12:97-136.

Verleger R, Jaśkowski P, Wascher E (2005) Evidence for an integrative role of P3b in linking reaction to perception. J Psychophysiol 19:165-181.

Walker KM, Bizley JK, King AJ, Schnupp JW (2011) Multiplexed and robust representations of sound features in auditory cortex. The Journal of neuroscience : the official journal of the Society for Neuroscience 31:14565-14576.

Winkler I, Czigler I, Sussman E, Horváth J, Balázs L (2005) Preattentive binding of auditory and visual stimulus features. J Cogn Neurosci 17:320-339.

This study was supported by the Russian Foundation for Humanities, research project No 15-06-10742.

Findings presented here were recently published: Chernyshev B. V., Bryzgalov D. V., Lazarev I. E., Chernysheva E. G. Distributed feature binding in the auditory modality: experimental evidence toward reconciliation of opposing views on the basis of mismatch negativity and behavioral measures // NeuroReport. - 2016. - T. 27, № 11. - C. 837-842.