

---

## BRIEF REPORTS

---

# The involuntary capture of attention by novel feature pairings: A study of voice–location integration in auditory sensory memory

FABRICE B. R. PARMENTIER

University of the Balearic Islands, Palma, Spain,  
University of Western Australia, Perth, Australia,  
and University of Plymouth, Plymouth, England

MURRAY T. MAYBERY

University of Western Australia, Perth, Australia

AND

JANE ELSLEY

University of Plymouth, Plymouth, England

Past researchers of the integration of information in memory have typically required participants to attend to and/or commit to memory the stimuli conveying distinct features, rendering difficult the examination of whether the maintenance of the feature pairings can occur involuntarily. To address this issue, the integration of voice and location information in auditory sensory memory was measured using a cross-modal oddball task, in which task-irrelevant auditory deviants are known to capture attention in an involuntary fashion. Participants categorized visual digits presented shortly after to-be-ignored sounds. These sounds consisted in the same phoneme played simultaneously in both ears but in different voices (female in one ear, male in the other). On most trials, the pairing of voice to location was constant (standard sound). On rare and unpredictable trials, the voices swapped locations (deviant sound). In line with past work on attention capture by auditory novelty, the participants were significantly slower to judge the visual digits following the deviant sound, indicating the involuntary encoding of the links between voice and location in auditory memory. These results suggest that voices and locations are integrated in memory and that this binding occurs in conditions in which participants do not intend to commit any information to memory.

---

The distinction between “what” and “where,” well established in the visual literature (e.g., Mishkin, Ungerleider, & Macko, 1983), is evidenced in the auditory modality by the finding of distinct neural pathways for a sound’s identity and location (e.g., Ahveninen et al., 2006; Arnott, Binns, Grady, & Alain, 2004; Rämä et al., 2004; Rauschecker & Tian, 2000; Romanski et al., 1999). Evidence of separate information processing for distinct auditory features has also been reported in behavioral studies. For example, it has been shown that memory for pitch and memory for location are selectively interfered with by frequency and location distractors, respectively (e.g., Anourova et al., 1999). Furthermore, distinct decay functions have been reported for pitch and loudness (Clément, Demany, & Semal, 1999). These findings indicate that, at least at some level, auditory features are processed and represented in-

dependently. Stimuli around us are not experienced as collections of independent features, however, and a challenge for scientists has therefore been to examine under what conditions distinct features are integrated within the brain (the *binding problem*).

Studies indicating that auditory features are integrated are relatively scarce and involve recognition paradigms. For example, participants performing a *same/different* decision on two consecutive sounds based on either their pitch or their loudness respond more accurately and faster if the task-irrelevant feature (loudness or pitch, respectively) is maintained constant across the two stimuli (Van Valkenburg et al., 2009). Such findings parallel earlier work in which it was reported that spoken words are better recognized if presented in the same voice at study and test (Bradlow, Nygaard, & Pisoni, 1999; Craik & Kirsner, 1974; Goh, 2005).

---

F. B. R. Parmentier, [fabrice.parmentier@uib.es](mailto:fabrice.parmentier@uib.es)

---

Whether binding in memory is contingent on voluntary memory encoding and the active deployment of attentional resources to support retention has recently become an issue of theoretical importance. So far, however, clear conclusions cannot be drawn. Furthermore, past attempts to examine this issue have been focused almost entirely on visual stimuli. The involvement of frontal cortical regions in the maintenance of visually presented verbal–spatial (Prabhakaran, Narayanan, Zhao, & Gabrieli, 2000) and object–location (Mitchell, Johnson, Raye, Mather, & D’Esposito, 2000) conjunctions suggests that some form of top-down control is required to prevent integrated representations from degrading (Wheeler & Treisman, 2002). Elsley and Parmentier (2009b) recently reported the abolition of binding between consonants and their locations in a probe recognition paradigm when attentional resources were depleted by the imposition of a concurrent memory load, a finding echoed with shapes and locations (Elsley & Parmentier, 2009a). These results suggest that the maintenance of integrated information may depend on the extent to which effort can be devoted to the encoding and retention of those features. Finally, some researchers have reported an associative asymmetry in the binding of visuospatial (Elsley, Parmentier, & Maybery, 2009), as well as auditory verbal–spatial (Maybery et al., 2009), features, indicating that attending to a single feature can result in the obligatory binding of a second feature to it, whereas the reverse is not observed, but that bidirectional associations (although not necessarily of equal strength) are observed when participants attend to both features.

However, other researchers have questioned the role of attention in the maintenance of bound representations. For example, Allen, Baddeley, and Hitch (2006) investigated the binding of visual features and reported no selective impact of concurrent tasks such as backward counting (see also Gajewski & Brockmole, 2006; Johnson, Hollingworth, & Luck, 2008). There is also evidence that the features of a visual object, such as its color and shape, are associated with its location, even in circumstances in which location is irrelevant to the recognition task at hand (Jiang, Olson, & Chun, 2000; Olson & Marshuetz, 2005).

To our knowledge, Maybery et al. (2009) are unique in examining the role of encoding demands on the maintenance of combined auditory features. Using a probe recognition paradigm, these authors observed binding between words and locations, as well as between tones and locations, when both features were encoded (even in conditions in which only one feature was judged). When only one feature was attended, however, spatial recognition judgments exhibited an advantage for probes involving the stimulus originally presented in that location, whereas no such advantage was observed for verbal recognition judgments, suggesting that attention to both features is sometimes necessary to observe binding.

One obstacle to the study of involuntary binding in memory is the use—traditionally—of tasks in which participants are often explicitly instructed to remember conjunctions of features and that invariably require them to voluntarily commit stimuli to memory. It is therefore difficult to ascertain whether binding can occur involuntarily

or whether it stems from the participants’ voluntary encoding and maintenance of the stimuli. In studies in which the researchers attempted to deplete the attentional resources available for the maintenance of bound representations by the use of secondary tasks, the interpretation of the results is complicated by the difficulty of quantifying the amount of attention used for the various task components and in establishing whether primary and secondary tasks truly draw from the same pool of resources. A strong test of involuntary binding of “what” and “where” in auditory memory would therefore be one in which a task that does not require participants to memorize any information at all is used. To this end, we measured binding indirectly, using a well-established paradigm capitalizing on the brain’s automatic detection of auditory novelty and the resulting measurable behavioral distraction in an unrelated task. We used the cross-modal oddball paradigm, adapted from electrophysiological studies (e.g., Escera, Yago, & Alho, 2001) and recently applied to the cognitive analysis of auditory distraction (Andrés, Parmentier, & Escera, 2006; Parmentier, 2008; Parmentier, Elford, Escera, Andrés, & San Miguel, 2008).

In this cross-modal oddball paradigm, participants categorize the parity of visual digits presented sequentially. A task-irrelevant sound is presented shortly (e.g., 300 msec) before each digit, which the participants are instructed to ignore. On most trials, the same sound is used (*standard*). On rare and random occasions, the standard is replaced by another sound (*deviant*). Electrophysiological studies indicate that deviants trigger three specific brain responses (MMN, P3a, and RON). MMN marks the detection of a change in a sequence of stimuli and reflects the result of a comparison between a memory trace for the standard sound and the current auditory signal (Näätänen & Winkler, 1999), thereby providing a tool to examine the nature of auditory sensory memory (e.g., Cowan, Winkler, Teder, & Näätänen, 1993; Winkler & Cowan, 2005). This first response triggers the preattentive orientation of attention toward a deviant sound, marked by the P3a response (Friedman, Cycowicz, & Gaeta, 2001). A later additional deflection (RON) is also observed in experiments in which participants engage in a primary task and is interpreted as the reorientation of attention toward the latter (e.g., Berti, Roeber, & Schröger, 2004). For the purpose of studying involuntary binding in auditory sensory memory, a distinct advantage of the cross-modal oddball task over the purely auditory oddball paradigm (e.g., Schröger, 1996) is that distractors and targets are temporally decoupled and presented in distinct sensory modalities, so that sounds are completely task irrelevant.

The distraction elicited by deviant sounds in the cross-modal oddball paradigm is measurable behaviorally as the difference in performance, in the visual task, between the deviant and standard conditions. Past studies have indeed shown that participants take significantly longer to categorize visual digits as odd or even when the digits are preceded by a novel or deviant sound relative to a standard (e.g., Andrés et al., 2006; Escera et al., 2001), an effect reflecting the time penalty resulting from involuntary switches of attention to and from the novel sound

(Parmentier et al., 2008). The critical feature of this paradigm is that distraction occurs only if the cognitive system detects a change between the current task-irrelevant sound stimulus and a memory representation of the standard. If standard and deviant sounds are indistinguishable, no distraction occurs (whether measured behaviorally or electrophysiologically).

In the present study, we used a modified version of the cross-modal oddball task. Specifically, we used standard and deviant sounds that differed only in terms of the pairings between “what” and “where” features to test whether these pairings form an intrinsic part of the memory representation for the standard sound. Our standard sound consisted of the utterance /a/ presented simultaneously in a male voice to one ear and in a female voice to the other ear. The deviant sound contained all of the features of the standard (the same locations and sounds) but in a re-paired arrangement, so that the male and female voices swapped locations. The hypothesis was simple: If the brain represents information in auditory memory in terms of free-floating traces only (independent traces hypothesis), the standard and deviant sounds should be indistinguishable, because they contain the same features (same locations, same sounds). Consequently, no change should be detected, and no distraction should be observed. However, if information in auditory sensory memory is represented as integrated sound–location objects (integration hypothesis), the deviant should be perceived as different from the standard, and distraction should be observed in the digit categorization task following the well-demonstrated capture of attention by auditory novelty (Andrés et al., 2006; Parmentier, 2008; Parmentier et al., 2008).

## METHOD

### Participants

Twenty-eight undergraduate psychology students (mean age = 19.97 years,  $SD = 4$  years; 18 female, 10 male) from the University of Plymouth took part in the experiment in exchange for course credit. All of the participants reported normal hearing and normal or corrected-to-normal vision.

### Materials

The sounds were built from recordings of the utterance /a/ in two voices, one male and one female. These 44-KHz, 16-bit mono recordings were digitally edited to 200-msec durations and normalized. Using these two sounds, two 200-msec stereo sound files were generated, in which the male and female recordings were presented simultaneously in the left and right channels, respectively (first sound), or vice versa (second sound). Across participants, each sound was used equally often as the standard and the deviant. The task was programmed in E-Prime and executed on a computer equipped with a 17-in. screen. The sounds were presented at approximately 70 dB through headphones.

### Design and Procedure

The participants were presented with three blocks of 336 trials each. A fixation cross was displayed at the center of the screen for the duration of the trial, except during the presentation of a visual digit. In each trial, a 200-msec sound was presented. Next, 100 msec after the sound's offset, a digit sustaining a viewing angle of approximately 2.6° appeared at the center of the screen for 200 msec, after which only the fixation cross was visible for another 700 msec for

the participants to categorize the digit as odd or even and to respond accordingly. The participants used the “Z” and “X” keys on the computer keyboard to respond, using two fingers of their dominant hand. The mapping between the response keys and the odd–even responses was counterbalanced across participants. The digits (1–8) were presented in random order (different for every participant), but with equal probabilities in each block.

Two sound conditions were compared within each block. In the standard condition (80% of trials), the sound consisted of the utterance /a/ presented in a male voice in the left channel and simultaneously in a female voice in the right channel for half of the participants (and vice versa for the other half). In the deviant condition (20% of trials), the sound consisted of the opposite mapping. The deviant trials were always preceded by at least one standard trial. In each block, the first 16 trials, which only contained standards, were treated as warm-up trials and were not included in the data analysis. Across the 960 test trials, 32 deviant trials were used in each successive group of 160 trials to ensure a relatively even, yet unpredictable, distribution of deviants across trials. The order of presentation of the standard and deviant trials was otherwise random and different for each participant. The participants were instructed to ignore the sounds, to concentrate on the digits task, and to respond as quickly and accurately as possible.

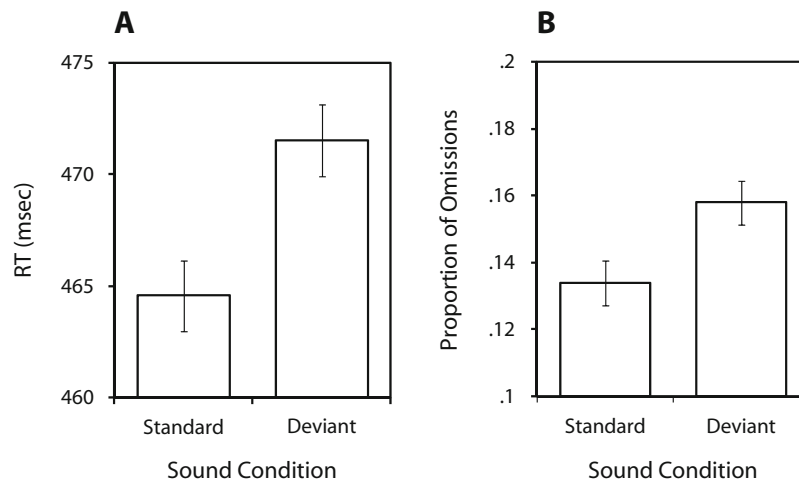
## RESULTS

Distraction was measured as the difference in performance on the categorization task between the deviant and standard conditions. Both response accuracy and response times (RTs) for correct responses were analyzed. The results were clear-cut, exhibiting significantly longer RTs in the deviant condition ( $M = 471.503$  msec,  $SD = 37.738$  msec) than in the standard condition ( $M = 463.557$  msec,  $SD = 33.237$  msec) [ $t(27) = 3.347$ ,  $p < .005$ ,  $d = .644$ ] (see Figure 1).

On a number of occasions, the participants failed to respond before the onset of the next trial. The analysis of the proportion of missing responses revealed a significantly higher number of these instances in the deviant condition ( $M = .158$ ,  $SD = .103$ ) than in the standard condition ( $M = .134$ ,  $SD = .075$ ) [ $t(27) = 2.935$ ,  $p < .01$ ,  $d = .565$ ]. When the participants responded, however, they did so with the high level of accuracy typically observed in this simple task and similarly in the deviant ( $M = .846$ ,  $SD = .062$ ) and standard ( $M = .845$ ,  $SD = .060$ ) conditions [ $t(27) < 1$ ,  $p = .944$ ,  $d = .014$ ].

## DISCUSSION

Our results are clear-cut: Evidence of the integration of voice and location information in auditory sensory memory was observed in a task in which the participants were not required to attend to the sounds, to commit them to memory, or to voluntarily maintain voice–location pairings in memory. This binding effect was measured as the distraction incurred in a visual, unrelated, primary task, in the form of a lengthening of response latencies and an increase in the number of late responses (in line with Parmentier et al.'s, 2008, demonstration that auditory deviants entail a time-consuming shift to and from the auditory deviant). The data therefore suggest that voice and location, despite the evidence indicating their independent process-



**Figure 1.** Mean response time (RT) and proportion of omissions in the standard and deviant conditions. Error bars represent the 95% confidence interval.

ing in the brain (e.g., Arnott et al., 2004), are integrated in memory and that this integration occurs without the participants' intention to maintain bound representations in memory. Sound appears to be associated with spatial location involuntarily, consistent with the idea that some forms of binding can occur preattentively, at least at a very basic level of processing (preattentive object files: Wolfe & Bennett, 1997). Our findings suggest that auditory sensory memory holds a representation of auditory events defined by what–where bindings instead of independent free-floating features (since our standard and deviant stimuli shared the same features and differed only in terms of the bindings between voice and location).

Although our study is unique in that we used behavioral distraction in a cross-modal oddball task to explore the nature of the binding in memory between voice and location, it is worth noting that our conclusion is in line with an electrophysiological study of binding between frequency and location (Winkler, Czigler, Sussman, Horváth, & Balázs, 2005). Winkler et al. used an auditory oddball task in which participants were presented with sequences of sounds varying in frequency and location. The participants were exposed to a presequence, in which most, but not all, combinations of frequencies and locations (both selected from fixed sets) were presented as standard tones. The goal of the presequence was to allow the building of a mental model of the standard tones in the participants' memory. A main sequence was then played, in which some of the combinations missing in the presequence were presented (deviants). The rationale was that since the features (frequency and location) making up these deviants were familiar to the participants, the observation of an MMN response would indicate that the memory representations for the standards included the associations between frequency and location. Such an MMN response was indeed observed, in line with previous studies in which the preattentive processing of irregular conjunctions of auditory features was examined (e.g., Sussman, Gomes, Nousak,

Ritter, & Vaughan, 1998). Our results indicate that the involuntary binding of auditory features is not limited to frequency and location features of simple tones (Winkler et al., 2005) or timbre and pitch (Takegata et al., 2005) and bring two novel findings: that involuntary binding also applies to human voice and location and that it can be reliably measured behaviorally (see Parmentier, 2008, for a model relating electrophysiology and behavior). It should be noted, finally, that evidence for the involuntary binding of features in auditory sensory memory does not preclude the independent processing of individual features. Indeed, several MMN studies have established the additive impact of independent deviants embedded in the same auditory events. For example, Schröger (1996) compared deviants defined by interaural time differences, interaural intensity differences, or both. Of interest, the combined deviant yielded an MMN signal very similar to the sum of the MMN responses measured in response to single feature deviants, indicating the parallel processing of interaural time and intensity differences and their additive effect on MMN (see also Wolff & Schröger, 2001). Hence, the evidence above highlights the flexibility of the cognitive system in detecting changes in individual feature channels, as well as changes relating to the relationship between features.

The key characteristic of our task, relative to other studies of binding in memory, is the demonstration of binding in a task that did not require participants to commit any information to memory. The participants had no reason to voluntarily learn the relationship between sounds and locations and had nothing to gain from it, yet the distraction yielded by a rare remapping of these features demonstrates that the neural model of past acoustic events includes the association between these features. This finding is in line with the results of a recent study by Zmigrod and Hommel (2009), in which participants responded to a first tone with a preestablished response and then to a subsequent tone on the basis of its pitch or loudness (Experiment 1), or

pitch or location (Experiment 2). Of particular relevance, Zmigrod and Hommel found that the benefit of repeating the relevant feature across the two sounds was greater when the irrelevant feature was also repeated, and that the cost of the alternation of the relevant feature across the sounds was reduced when the relevant feature was also alternated. These findings also suggest that auditory features are integrated in memory even when the task does not require committing information to memory and are in line with the proposition that features forming part of an event are integrated and maintained in event files (Hommel, 2004; see, e.g., Roeber, Berti, Widmann, & Schröger, 2005, for evidence of the role of response codes in novelty distraction). It is worth pointing out that our study differs from Zmigrod and Hommel's insofar as theirs required participants to attend to the auditory stimuli. This implies that the binding that they demonstrated, if unintentional insofar as the participants had no reason to commit the sounds to memory, cannot be argued to be independent of whether the sounds were attended. In our task, in contrast, participants were instructed to ignore the sounds, which offered no help in categorizing unrelated visual targets. In summary, auditory features appear to be integrated in memory in an involuntary fashion, whether participants are required to ignore sounds (the present task) or to attend to them (Zmigrod & Hommel, 2009).

Although the cross-modal oddball task provides a way of measuring the involuntary binding of auditory features when participants are instructed to ignore the sounds, further research is necessary to establish whether our findings generalize to features other than voice and location. In particular, the processing of voice information may require further investigation, given the evidence suggesting that voice features enjoy a special status among other auditory signals (Fifer & Moon, 1994; Levy, Granot, & Bentin, 2001; Rauschecker, Tian, & Hauser, 1995). It remains therefore possible that the behavioral impact of the distraction yielded by the deviants in our oddball task reflects, in part, binding processes that are specific to voices and locations.

Finally, the findings reported here also have implications for our understanding of auditory sensory memory with which new upcoming stimuli are compared and that form the basis of the MMN response (e.g., Cowan et al., 1993). The functional characteristics of auditory sensory memory are still relatively poorly understood, although it has been demonstrated that it can hold a representation of the standard for several seconds (Winkler & Cowan, 2005) and does not seem to be subject to interference as is short-term memory (Winkler, Cowan, Csépe, Czigler, & Näätänen, 1996). Future researchers will undoubtedly seek to address important questions, such as the capacity of auditory sensory memory, and the corollary issue of the nature of the representations maintained by such memory. The present study suggests that auditory sensory memory holds not only the constituent features of past acoustic events, but also their association. This association was, in our study, between a feature defining the sound's identity (voice) and a more transient one that does not form an intrinsic part of the auditory object per se (location).

## AUTHOR NOTE

This research was supported by ESRC Grant RES-062-23-0241 to F.B.R.P. and Australian Research Council Grants DP0452312 and DP0773836 to M.T.M., F.B.R.P., and D. Jones. Correspondence concerning this article should be addressed to F. B. R. Parmentier, Department of Psychology, University of the Balearic Islands, Ctra. de Valldemossa, km 7.5, 07122 Palma, Spain (e-mail: fabrice.parmentier@uib.es).

## REFERENCES

- AHVENINEN, J., JÄÄSKELÄINEN, I. P., RAIJ, T., BONMASSAR, G., DEVORE, S., HÄMÄLÄINEN, M., ET AL. (2006). Task-modulated "what" and "where" pathways in human auditory cortex. *Proceedings of the National Academy of Sciences*, **103**, 14608-14613.
- ALLEN, R. J., BADDELEY, A. D., & HITCH, G. J. (2006). Is the binding of visual features in working memory resource-demanding? *Journal of Experimental Psychology: General*, **135**, 298-313.
- ANDRÉS, P., PARMENTIER, F. B. R., & ESCERA, C. (2006). The effect of age on involuntary capture of attention by irrelevant sounds: A test of the frontal hypothesis of aging. *Neuropsychologia*, **44**, 2564-2568.
- ANOUROVA, I., RÄMÄ, P., ALHO, K., KOIVUSALO, S., KALMARI, J., & CARLSON, S. (1999). Selective interference reveals dissociation between auditory memory for location and pitch. *NeuroReport*, **10**, 3543-3547.
- ARNOTT, S. R., BINNS, M. A., GRADY, C. L., & ALAIN, C. (2004). Assessing the auditory dual-pathway model in humans. *NeuroImage*, **22**, 401-408.
- BERTI, S., ROEBER, U., & SCHRÖGER, E. (2004). Bottom-up influences on working memory: Behavioral and electrophysiological distraction varies with distractor strength. *Experimental Psychology*, **51**, 249-257.
- BRADLOW, A. R., NYGAARD, L. C., & PISONI, D. B. (1999). Effects of talker, rate, and amplitude variation on recognition memory for spoken word. *Perception & Psychophysics*, **61**, 206-219.
- CLÉMENT, S., DEMANY, L., & SEMAL, C. (1999). Memory for pitch versus memory for loudness. *Journal of the Acoustic Society of America*, **106**, 2805-2811.
- COWAN, N., WINKLER, I., TEDER, W., & NÄÄTÄNEN, R. (1993). Short- and long-term prerequisites of the mismatch negativity in the auditory event-related potential (ERP). *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **19**, 909-921.
- CRAIK, F. I. M., & KIRSNER, K. (1974). The effect of speaker's voice on word recognition. *Quarterly Journal of Experimental Psychology*, **26**, 274-284.
- ELSLEY, J. V., & PARMENTIER, F. B. R. (2009a). *Is binding in visuospatial working memory impaired by a concurrent memory load?* Manuscript submitted for publication.
- ELSLEY, J. V., & PARMENTIER, F. B. R. (2009b). *Is verbal-spatial binding in working memory impaired by a concurrent memory load?* *Quarterly Journal of Experimental Psychology*, **62**, 1696-1705.
- ELSLEY, J. V., PARMENTIER, F. B. R., & MAYBERRY, M. T. (2009). *Feature binding within visuo-spatial working memory: Asymmetry between shape and location encoding.* Manuscript submitted for publication.
- ESCERA, C., YAGO, E., & ALHO, K. (2001). Electrical responses reveal the temporal dynamics of brain events during involuntary attention switching. *European Journal of Neuroscience*, **14**, 877-883.
- FIFER, W. P., & MOON, C. M. (1994). The role of mother's voice in the organization of brain functions in the newborn. *Acta Paediatrica*, **83**, 86-93.
- FRIEDMAN, D., CYCOWICZ, Y. M., & GAETA, H. (2001). The novelty P3: An event-related brain potential (ERP) sign of the brain's evaluation of novelty. *Neuroscience & Biobehavioral Reviews*, **25**, 355-373.
- GAJEWSKI, D. A., & BROCKMOLE, J. R. (2006). Feature bindings endure without attention: Evidence from an explicit recall task. *Psychonomic Bulletin & Review*, **13**, 581-587.
- GOH, W. D. (2005). Talker variability and recognition memory: Instance-specific and voice-specific effects. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **31**, 40-53.
- HOMMEL, B. (2004). Event files: Feature binding in and across perception and action. *Trends in Cognitive Sciences*, **8**, 494-500.
- JIANG, Y., OLSON, I. R., & CHUN, M. M. (2000). Organization of visual short-term memory. *Journal of Experimental Psychology: Learning, Memory & Cognition*, **26**, 683-702.

- JOHNSON, J. S., HOLLINGWORTH, A., & LUCK, S. J. (2008). The role of attention in the maintenance of feature bindings in visual short-term memory. *Journal of Experimental Psychology: Human Perception & Performance*, **34**, 41-55.
- LEVY, D. A., GRANOT, R., & BENTIN, S. (2001). Processing specificity for human voice stimuli: Electrophysiological evidence. *NeuroReport*, **12**, 2653-2657.
- MAYBERY, M. T., CLISSA, P. J., PARMENTIER, F. B. R., LEUNG, D., HARSA, G., FOX, A. M., & JONES, D. M. (2009). Binding of verbal and spatial features in auditory working memory. *Journal of Memory & Language*, **61**, 112-133.
- MISHKIN, M., UNGERLEIDER, L. G., & MACKO, K. A. (1983). Object visuo and spatial vision: Two cortical pathways. *Trends in Neurosciences*, **6**, 414-417.
- MITCHELL, K. J., JOHNSON, M. K., RAYE, C. L., MATHER, M., & D'ESPOSITO, M. (2000). Aging and reflective processes of working memory: Binding and test load deficits. *Psychology & Aging*, **15**, 527-541.
- NÄÄTÄNEN, R., & WINKLER, I. (1999). The concept of auditory stimulus representation in cognitive neuroscience. *Psychological Bulletin*, **125**, 826-859.
- OLSON, I. R., & MARSHUETZ, C. (2005). Remembering "what" brings along "where" in visual working memory. *Perception & Psychophysics*, **67**, 185-194.
- PARMENTIER, F. B. R. (2008). Toward a cognitive model of distraction by auditory novelty: The role of involuntary attention capture and semantic processing. *Cognition*, **109**, 345-362.
- PARMENTIER, F. B. R., ELDFORD, G., ESCERA, C., ANDRÉS, P., & SAN MIGUEL, I. (2008). The cognitive locus of distraction by acoustic novelty in the cross-modal oddball task. *Cognition*, **106**, 408-432.
- PRABHAKARAN, V., NARAYANAN, K., ZHAO, Z., & GABRIELI, J. D. E. (2000). Integration of diverse information in working memory within the frontal lobe. *Nature Neuroscience*, **3**, 85-90.
- RÄMÄ, P., POREMBA, A., SALA, J. B., YEE, L., MALLOY, M., MISHKIN, M., & COURTNEY, S. M. (2004). Dissociable functional cortical typologies for working memory maintenance of voice identity and location. *Cerebral Cortex*, **14**, 768-780.
- RAUSCHECKER, J. P., & TIAN, B. (2000). Mechanisms and streams for processing of "what" and "where" in auditory cortex. *Proceedings of the National Academy of Sciences*, **97**, 11800-11806.
- RAUSCHECKER, J. P., TIAN, B., & HAUSER, M. (1995). Processing of complex sounds in the macaque nonprimary auditory cortex. *Science*, **268**, 111-114.
- ROEBER, U., BERTI, S., WIDMANN, A., & SCHRÖGER, E. (2005). Response repetition vs. response change modulates behavioral and electrophysiological effects of distraction. *Cognitive Brain Research*, **22**, 451-456.
- ROMANSKI, L. M., TIAN, B., FRITZ, J., MISHKIN, M., GOLDMAN-RAKIC, P. S., & RAUSCHECKER, J. P. (1999). Dual streams of auditory afferents target multiple domains in the primate prefrontal cortex. *Nature Neuroscience*, **2**, 1131-1136.
- SCHRÖGER, E. (1996). Interaural time and level differences: Integrated or separated processing? *Hearing Research*, **96**, 191-198.
- SUSSMAN, E., GOMES, H., NOUSAK, J. M., RITTER, W., & VAUGHAN, H. G., JR. (1998). Feature conjunctions and auditory sensory memory. *Brain Research*, **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. *Cognitive Brain Research*, **25**, 169-179.
- VAN VALKENBURG, D., MAYBERY, M. T., LEUNG, D., KUBOVY, M., PARMENTIER, F. B. R., & JONES, D. M. (2009). *The binding of pitch and loudness features in auditory memory*. Manuscript submitted for publication.
- WHEELER, M. E., & TREISMAN, A. M. (2002). Binding in short-term visual memory. *Journal of Experimental Psychology: General*, **131**, 48-64.
- WINKLER, I., & COWAN, N. (2005). From sensory to long-term memory. Evidence from auditory memory reactivation studies. *Experimental Psychology*, **52**, 3-20.
- WINKLER, I., COWAN, N., CSÉPE, V., CZIGLER, I., & NÄÄTÄNEN, R. (1996). Interactions between transient and long-term auditory memory as reflected by the mismatch negativity. *Journal of Cognitive Neuroscience*, **8**, 403-415.
- WINKLER, I., CZIGLER, I., SUSSMAN, E., HORVÁTH, J., & BALÁZS, L. (2005). Preattentive binding of auditory and visual stimulus features. *Journal of Cognitive Neuroscience*, **17**, 320-339.
- WOLFE, J. M., & BENNETT, S. C. (1997). Preattentive object files: Shapeless bundles of basic features. *Vision Research*, **37**, 25-43.
- WOLFF, C., & SCHRÖGER, E. (2001). Human pre-attentive auditory change-detection with single, double, and triple deviations as revealed by mismatch negativity additivity. *Neuroscience Letters*, **311**, 37-40.
- ZMIGROD, S., & HOMMEL, B. (2009). Auditory event file: Integrating auditory perception and action planning. *Attention, Perception, & Psychophysics*, **71**, 352-362.

(Manuscript received May 5, 2009;  
revision accepted for publication September 11, 2009.)