Title
Enhanced Performance for Recognition of Irrelevant Target-Aligned Auditory Stimuli: Unimodal and Cross-modal Considerations

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 34(34)

ISSN
1069-7977

Authors
Dewald, Andrew
Sinnett, Scott

Publication Date
2012

Peer reviewed

Enhanced Performance for Recognition of Irrelevant Target-Aligned Auditory Stimuli: Unimodal and Cross-modal Considerations

Andrew D. Dewald (adewald@hawaii.edu)
Department of Psychology, University of Hawaii at Manoa
2530 Dole Street, Honolulu, HI 96822 USA

Scott Sinnett (ssinnett@hawaii.edu)
Department of Psychology, University of Hawaii at Manoa
2530 Dole Street, Honolulu, HI 96822

Abstract

Task-irrelevant stimuli are later recognized at enhanced levels providing that they had previously appeared with a task-relevant target (Dewald & Sinnett, submitted; Seitz & Watanabe, 2003, 2005). The present investigation explores this notion in the auditory sensory modality. Participants listened to a stream of auditory sounds and spoken words with the instruction to detect repetitions in only the sound stream (i.e., ignore the words). A surprise test measured recognition for the previously played words. Overall, when comparing target-aligned and non-aligned information in a later recognition task, facilitation was observed for words that had been aligned with target repetitions, despite equal presentation frequency and being irrelevant to the primary repetition task. This enhancement was mediated by the sensory modality of presentation in the surprise recognition task. Congruent auditory presentations between the exposure and recognition tasks yielded improved performance, and under cross-modal presentations the magnitude of the enhancement was greatest.

Introduction

An emerging body of research has explored how unattended information (explicitly or implicitly presented) that appears simultaneously with a task-target in a separate task is later recognized in a subsequently presented surprise recognition test (Dewald, Sinnett, & Doumas, 2011; Seitz & Watanabe, 2003, 2005; Swallow & Jiang, 2010, 2011; Tsushima, Sasaki, & Watanabe, 2006; Tsushima, Seitz, & Watanabe, 2008). These investigations have generally consisted of a primary task involving visually presented task-relevant and previously irrelevant stimuli, with a later surprise recognition test for the irrelevant stimuli. Collectively, the findings suggest that the task-irrelevant stimuli are indeed processed, as long as they were previously aligned with a task-relevant target. However, two opposing patterns have been observed, with facilitatory (see for example Seitz & Watanabe, 2003; Watanabe, Nañez, & Seitz, 2001) or inhibitory (see for example Tsushima et al., 2006, 2008) effects for the previously aligned, but irrelevant material, seemingly dependent on whether it had originally been presented below or above, respectively, the threshold for conscious awareness.

Seitz and Watanabe (2003) required participants to identify differently colored target letters in a rapid serial visual presentation (RSVP) of the letters. Importantly, an array of randomly moving dots (irrelevant to the task) was presented in the background during the letter identification task, of which a small subset moved coherently. Despite the coherent motion being implicit (i.e., participants were unable to reliably detect the coherent motion), participants were proficient at discriminating the motion (still presented below threshold) in a later discrimination task as long as it had previously been exposed simultaneously with the presence of a differently colored target letter during the identification task.

It is important to note that the initial motion presentation was implicit. This is a key point, as when presenting the coherent motion above threshold (i.e., 50% of the dots moved coherently), Tsushima et al (2008) showed an inhibition for target-aligned motions. Thus, it appears that facilitatory and inhibitory effects might be dependent on whether the initial presentation was presented below or above threshold, respectively.

As the majority of this research has used a relatively non-complex and simple stimuli (e.g., coherent motion in an array of moving dots), we recently examined (see Dewald et al., 2011) how a more salient stimulus, but irrelevant to the primary task, might affect later processing during an adapted inattentional blindness (IB) task (see for example Rees, Ruth, Frith, & Driver, 1999; Sinnett, Costa, & Soto-Faraco, 2006). Participants were required to monitor a stream of pictures that had written words superimposed on top of each image, and respond to immediate repetitions in the picture steam while ignoring the word stream. A subsequently presented surprise recognition test for the previously ignored words demonstrated that words that had been temporally aligned with the presence of a task-target (i.e., an immediately repeating picture) were recognized significantly below chance levels (i.e., inhibited), while words that had been temporally aligned with non-targets (i.e., a non repeating picture) were recognized at chance levels.

It is important to note that the highly salient stimuli (written words) were presented for 350 ms (i.e., more than enough time for explicit perception). However, the aligned words were nonetheless inhibited in a later surprise recognition test. This finding dovetails with the earlier described findings by Tsushima et al (2008), who used target-aligned coherent motion as a stimulus and observed an analogous inhibition of motion discrimination for target-aligned motions. That is, the coherent motion was also presented above threshold.

Other experimental paradigms have utilized different approaches to further investigate the way that the temporal alignment of task relevant and irrelevant information affects the later recognition of the irrelevant stimuli. Interestingly, despite the explicit presentation of their stimuli, the opposite findings, have been demonstrated, with a facilitation being observed for seemingly analogous conditions that elicited an inhibition. For instance, Swallow and Jiang (2010; see also Lin et al, 2010) completed a series of experiments suggesting an "attentional boost" (i.e., facilitation) for simultaneously presented, suprathreshold information in a dual-task paradigm. Participants monitored a stream of pictures of various scenes, while a series of items (small black superimposed shapes) was simultaneously paired with the presentation of each picture. The task was to remember as many of the presented scenes as possible, in addition to monitor the distractor stream of shapes for the presence of a color change. In a subsequent forced choice recognition test for the picture scenes, an enhanced recognition for pictures that had been previously presented simultaneously with the presence of a target (i.e., differently colored shape) in the distractor stream was observed.

Of particular note to Swallow and Jiang's (2010) findings is that participants were required to attend to *both* streams of information simultaneously. Recall that in the paradigms utilized by Dewald et al (2011) and Tsushima et al (2006, 2008), participants were instructed to detect a target in one stream, but explicitly instructed not to attend to the other stream of irrelevant information, which constituted the information to be recognized in the subsequent surprise test. The division of attention between both streams of information in Swallow and Jiang's paradigm could be the reason why a facilitation was observed in their results, rather than an inhibition.

A recent investigation by Dewald and Sinnett (submitted) aimed to more closely align their paradigm with the original procedure used by Tsushima et al (2006; see also, Seitz & Watanabe, 2003; Watanabe et al, 2001). In their experiment the exact same motion was paired with all of the target letters. This is different from the paradigm used by Swallow and Jiang (2010) or Dewald et al (2011), where a number of different pictures or words (irrelevant items in the primary task, but the items of interest in the subsequent recognition test), respectively, were aligned with targets in the primary task. In fact, the initial paradigm used by Dewald et al (2011) had 50 different words serving as the irrelevant stimuli during the picture repetition detection task. Therefore, in our subsequent work (Dewlad & Sinnett, submitted) this relatively infrequent exposure rate (each of the 50 aligned words was presented only four times) was increased to an exposure rate more similar to Tsushima et al (2006, 2008). Accordingly, only a single word was aligned with all of the targets from the primary task of detecting picture repetitions. This enabled the number of instances that this task-irrelevant word was presented to be greatly increased (from two times per word per participant, to 120 times). If the premise is that explicit information is later inhibited during a recognition task if it had been aligned with a task-relevant target, then an inhibition should have been observed for these items, especially given that the paradigm better replicated the original work demonstrating such an inhibition by Tsushima et al (2006). However, a facilitation for aligned words was seen, suggesting that the relationship between inhibition and facilitation is more complex than just whether the previously aligned stimuli were explicitly or implicitly presented. It is likely possible that the salience of the stimuli is also important, as an ignored written word is arguably processed to a higher level (i.e., semantically) than the ignored coherent motion of a moving array of dots.

Assuming that the facilitatory effect for explicitly presented, but ignored, stimuli is driven by whether the irrelevant stimulus is temporally aligned with the presence of a task-relevant target, it is important to extend these results to other sensory modalities. Despite vision being the dominant sense in humans (Chandra, Robinson, & Sinnett, 2011; Colavita et al., 1974; Posner et al., 1980; Sinnett et al., 2007), it is clear that the human perceptual experience is a result of multisensory information. Thus, it is important to explore if these inhibitory and faciltatory effects extend to other sensory modalities, as this will further inform how information is processed both within, and across modalities. For instance, Sinnett et al. (2006) demonstrated that when attentional reservoirs were depleted by a primary task, inattentional blindness for spoken word perception was interrupted to the same degree as visual word recognition, however performance improved under multimodal conditions. Furthermore, a recent investigation by Dewald and Sinnett (2011a) extended this finding to the auditory modality by including an additional analysis for items that had previously appeared simultaneously with targets in the separate task. In this case, an inhibition for spoken words (explicitly presented) was observed. However, it should be noted that this investigation also used a large number of auditory words as irrelevant stimuli (i.e., identical to the visual paradigm incorporated in Dewald et al., 2011), therefore it is unknown if the inhibition for target-aligned stimuli will extend to conditions with an increased exposure rate (i.e., similar to the visual condition of Dewald & Sinnett, submitted; see also Tsushima et al., 2006, 2008).

Addressing precisely this, we adapted the same paradigm utilized in Dewald & Sinnett (submitted) to an auditory presentation, with the primary task to detect target repetitions in the sound stream, and the secondary task to subsequently recognize the previously ignored words that had been played simultaneously with the sounds. Critically, only one specific spoken word was presented to participants. This increased exposure rate lead to an enhanced performance rather than inhibited (i.e., akin to the attentional boost observed by Swallow & Jiang, 2010). If visual findings (Dewald & Sinnett, submitted) extend to the auditory modality, then we predict that the higher exposure rate of the irrelevant spoken word will lead to a later facilitation in recognition for that word, as long as it had appeared simultaneously with an attended target in the previous task[1].

A separate but equally important aspect of this research explores the modality in which the surprise test is presented. The use of dual task paradigms is pervasive in the cognitive sciences. For instance, seminal studies on dichotic listening (Broadbent, 1958; Treisman, 1964) presented participants with orthogonal messages to each ear, with the instruction to only direct their attention to a single channel of information. After this initial task, an unexpected test was given to assess the ability to recognize or recall the information that had been presented previously at the ignored ear. Of key importance here is the consideration of the sensory modality that the surprise test was presented in, respective of the initial exposure during the primary task. For instance, in these classic studies, the surprise test that probed participants' ability to process the originally presented irrelevant information was always given in the same modality as the original presentation during exposure in the primary task (auditorily).

To the best of our knowledge, this congruency in modality presentation between exposure and recognition tests has never been systematically manipulated. Therefore, given the recent interest in extending inattentional blindness (see Sinnett et al., 2006) and target-alignment findings (see Dewald & Sinnett, 2010, 2011a,b) to the auditory modality and across modalities, it is important to explore how presenting the surprise test in a congruent modality would affect results, if at all. Thus, in addition to the aforementioned goals of extending this paradigm to the auditory modality and incorporating a higher exposure rate, we also presented the surprise recognition test in the same or different sensory modality, or across modalities. If primary and secondary task modality congruence is a factor, then we would expect improved results for congruent matchings vs. incongruent matchings, and potentially an additional enhancement for multimodal presentations given that performance is generally enhanced for multisensory presentations (see Driver & Spence, 2004).

Method

***Participants*.** Fifty-one participants were recruited from the University of Hawai'i at Manoa in exchange for course credit. Each participant completed the same auditory repetition detection task, but were divided across three different types of surprise recognition tests: visual only (n=18), auditory only (n=17), or cross-modal (n=16). Participants were naïve to the experiment and had normal or corrected to normal hearing.

Materials. A total of 16 one to two syllable, high-frequency English words (average length of 5 letters, range of 4-6 letters) were selected from the MRC psycholinguistic database (Wilson, 1988). The overall average frequency of the 150 selected words was 120 per million, ranging between 28 and 686. A native English speaker's voice was recorded reading the list three times, after which three blind listeners chose the best exemplar of each spoken word (a fourth listener was recruited in order to break a tie when needed) The selected recordings were edited to have the same length of presentation (350 ms) and average amplitude. The sound stimuli were extracted from a database of 100 familiar sounds and were also edited to 350 ms and similar average amplitude (see Sinnett et al., 2006). A stream of 960 sound-word concatenated items was created. Repeated sounds acted as the task relevant-targets. The presentation stream was broken into eight blocks of 120 trials in which an immediate sound repetition occurred on average one out of every eight trials, equating to 15 task-relevant target repetitions per block.

Eight spoken words (of the original 16) were randomly selected to overlay the 960 trial sound stream. This was done to parallel the quantity of items and exposure to irrelevant stimuli (see Dewald & Sinnett, submitted; Tsushima et al., 2006) as well as the dependent measure employed by Dewald and Sinnett (submitted; i.e., the analogous experiment in the visual domain). The presentation was pseudorandomized so that on average one out of every eight trials was an immediate sound repetition (and therefore the presentation of the same superimposed task-irrelevant target word). Critically, only one superimposed spoken word was aligned with all of the immediately repeated sounds for each participant. This single word was randomized between the eight words between participants, so as to control for any possible differences that may have existed regarding the saliency of any particular word. Lastly, it is important to note that both aligned and non-aligned words were exposed to participants in equal amounts.

A surprise recognition test for the presented words was administered after the completion of the repetition detection task. The test consisted of a total of sixteen words from which half came from the previously heard sound stream, while the other half consisted of foil words that had never been heard before (average frequency of 236 per million with a range of 165-399). The word recognition tasks were randomized and presented by DMDX software (http://www.u.arixona.edu/jforster/dmdx.htm) one at a time, in either the visual or auditory modality, or across modalities. For the visual presentation the words were written in bold, capitalized letters in Arial font at a size of 24 points, and remained on the screen until a response was made. For auditory presentations the words were spoken just as they were in the initial repetition detection task, albeit without the accompanying sounds. The sound stream and relevant surprise tasks were presented from two external speakers, equidistant to the computer screen. Cross-modal presentations involved the written word on the screen with the spoken word presented simultaneously.

[1] Note, target aligned and non-aligned words were themselves exposed in equal proportions. The higher frequency relates to comparisons to previous, but analogous research (see Dewald et al., 2011; Dewald & Sinnett, 2011, submitted).

*Procedure***.** Participants were required to attend to the sound stream (i.e., they were explicitly instructed to ignore the simultaneously presented, overlaid spoken words) and respond to immediate sound repetitions by pressing the 'G' key on the keyboard of the computer. Each item in the sound-word presentation was presented for 350 ms with a 150-ms inter-stimulus interval (ISI; silence) for a stimulus onset asynchrony (SOA) of 500 ms. Before the first experimental block, a training block of eight trials was given and repeated until participants were familiar and comfortable with the task.

Immediately after the repetition detection task, the surprise word recognition test was administered to all participants (modality type of surprise task dependent on condition). Participants were instructed to press the "B" key if they had heard the word during the repetition detection task or, instead, the "V" key if they had not heard the word before.

Results

To assess whether recognition performance was modulated by target alignment or the modality of presentation of the surprise task, a two-way, repeated measures ANOVA was conducted, with surprise test modality (auditory, visual, or cross-modal) as a between-subjects factor, and target alignment (target-aligned or non-aligned) as a within subjects factor. A main effect for target alignment confirmed that, overall, word recognition performance was significantly better for *target-aligned* words (72.5%) when compared to *non-aligned* words (58.8.%) (F (2, 48) = 3.54, p = .03). No main effect was found for modality of presentation (F (2, 48) = 1.58, p = .211). Additionally, and of key importance to this analysis, an interaction was observed between target-alignment and modality of presentation (F (2, 48) = 3.08, $p < .05$), suggesting that the modality of presentation in the recognition task played a role in recognition performance between *target-aligned* and *non-aligned* words. To further explore this interaction, each surprise recognition condition (auditory, visual, or cross-modal) is individually analyzed below.

Visual surprise recognition test (VR): Overall task performance for the surprise test was 65.8%, which was statistically different from chance (t (46) = 3.08, p = .001). More importantly, the recognition for the *target-aligned* words (61.0%, SE=.11) was not statistically different from *non-aligned* words (61.5%, SE=.57; t (17)= .03, p = .97; see Figure 2). Note that the overall performance was higher due to increased performance on correctly rejecting foils. The correct rejection of foil words was compared with overall performance for *target-aligned* and *non-aligned* words. There was a significant difference between recognition for *target-aligned* words and correct rejections (*target-aligned*: 61.0%, SE=.11 vs. *CR*: 79.6%, SE=.04, t(17)=2.14, p = .02) as well as a significant difference between correctly recognizing *non-aligned* words and correct rejections (*non-aligned*: 61.5%, SE=.57 vs. *CR*: 79.6%, SE=.04, t(17)=2.47, p = .02). Lastly, confirming that participants were able to successfully perform the initial repetition task, the target repetitions were significantly detected (**Hits:** 68% *SE*=.18 vs. **Misses:** 28% *SE*=.04, t(17)=9.42, $p < .001$).

Auditory surprise recognition test (AR): Overall task performance was 71.0%, which was statistically different from chance (t (45) = 4.47, p = .001). Contrary to the visual condition, recognition for *target-aligned* words (76.4%, SE=.10) was significantly better than *non-aligned* words (58.8%, SE=.05, t (16)= 2.37, p = .05; see Figure 1). Again, unlike the first experiment, there was no difference in performance between *target-aligned* word recognition and correct rejections of foils (*target-aligned*: 76.04%, SE=.10 vs. *CR*: 77.7%, SE=.06, t(16)=0.09, p = .967) However, there was a significant difference between *non-aligned* word recognition and correctly rejecting foil words (*non-aligned*: 58.8%, SE=.05 vs. *CR*: 77.7%, SE=.06, t(16)=2.35, p = .03). Lastly, participants accurately detected target repetitions (**Hits:** 73% *SE*=.10 vs. **Misses:** 23% *SE*=.11, t(16)=8.47, $p < .001$).

Cross-modal surprise recognition test (CR). Overall word recognition was 69.2%, which was statistically better than chance (t (43) = 4.28, p = .001). Recognition for *target-aligned* words (81.2%, SE=.10) was significantly better than *non-aligned* words (55.7%, SE=.05; t (15)= 2.59, p = .05; see Figure 1). There was no difference in performance between *target-aligned* word recognition and correct rejections of foils (*target-aligned*: 81.2%, SE=.10 vs. *CR*: 70.8%, SE=.06, t(15)=0.82, p = .422). Again, however, there was a significant difference between *non-aligned* word recognition and correctly rejecting foil words (*non-aligned*: 58.7%, SE=.05 vs. *CR*: 70.8%, SE=.06, t(15)=2.07, p =.05). Additionally, participants accurately detected immediate sound target repetitions in the primary task (**Hits:** 75% *SE*=.13 vs. **Misses:** 24% *SE*=.03, t(15)=9.23, $p < .001$).

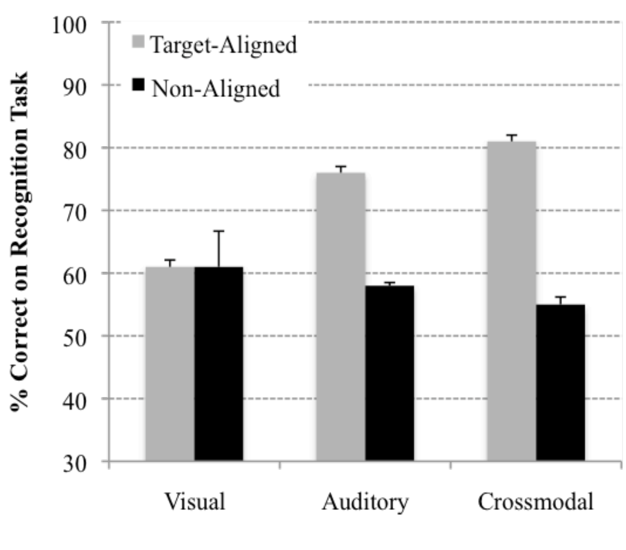

Figure 1. Recognition percentages for Target-Aligned (grey bar) and Non-Aligned (black bar) words dependent on the modality of the surprise test.

Discussion

The present findings extend investigations exploring how above threshold, but unattended information is processed when it appears simultaneously with an attended target. We demonstrate that in all conditions, the overall recognition of the words in the surprise task was better than chance, despite attention not being directed to the words. This finding is contrary to what can be predicted from inattentional blindness investigations in that the overall recognition of the words should have been at chance, or perhaps even inhibited for the irrelevant stimulus (Dewald et al, 2011; Rees et al, 1999; Sinnett et al, 2006). It is likely that the increased exposure to (i.e., fewer items) to the irrelevant stimuli drove this effect. More importantly, the interaction between target-alignment and the modality of the surprise test was significant, driven by performance for *target-aligned* words being statistically better than non-aligned words in the **auditory recognition (**76% vs. 58%) and **cross-modal recognition** (81% vs. 55%) conditions. Note, when comparing the magnitude of the enhancement between these two conditions, there was (no) additional improvement for cross-modal conditions. Accordingly, this suggests that, at least in the present case, temporally aligning explicitly presented, irrelevant, auditory stimuli with relevant auditory target stimuli facilitates subsequent recognition of the irrelevant stimuli, but only if the recognition test is presented in the same modality as the initial task, or across modalities. Thus, an "attentional boost" (see Swallow & Jiang, 2010) for irrelevant stimuli was observed, as long as they were initially presented simultaneously with a target in the picture repetition task, despite not receiving direct attention.

Interestingly, when the surprise recognition task was presented in an incongruent modality from the exposure (i.e., the visual surprise recognition task) there was no difference between *target-aligned* and *non-aligned* words. Previous findings have demonstrated an inhibition for both visually aligned (Dewald et al., 2011) and auditorily aligned words (Dewald & Sinnett, 2011a) in a dual-task paradigm, despite not controlling for task modality congruence (i.e., the visual example presented information in the congruent modality, while the auditory example also presented the surprise test in the visual modality). However, it should be noted that both of these examples used a much lower exposure rate (i.e., 50 words had been aligned in the repetition detection task rather than only one that was repeated 120 times). This could possibly explain why we failed to observe a difference in the visual task condition here.

Further complicating the matter, Dewald and Sinnett (submitted) observed an enhancement for target-aligned words under nearly isomorphic conditions as utilized here, with the only difference being that the initial repetition detection task was presented in the visual modality. In fact, all research conducted thus far, exploring the fate of irrelevant auditory words either target-aligned or not, has presented the words in the visual modality during the subsequent recognition task (Dewald & Sinnett, 2011a; Sinnett et al, 2006; see Dewald & Sinnett, 2011b for a cross-modal example). It is possible that the incongruence between the modality of presentation and subsequent recognition may in some way affect the recognition of the previously unattended items. Thus, further research should explore whether the observed enhancement for visually aligned words when tested with a visually presented surprise task extends to incongruent task modality presentations (or across modalities).

In the present experiment the cross-modal presentation lead to the greatest magnitude of enhancement for the previously aligned words in the surprise recognition test. This aligns well with previous investigations of attentional allocation across sensory modalities in perceptual and recognition tasks, suggesting that cross-modal presentations generally lead to superior to performance when compared to unimodal presentations (Dewald & Sinnett, 2011b; Duncan, Martens, & Ward, 1997; Sinnett et al, 2006; Toro, Sinnett, & Soto-Faraco, 2005).

The possibility that the presentation modality of the surprise test could modulate performance has yet to be fully explored. The resulting recognition performance, observed when the surprise task was incongruent to the initial exposure, necessitates an investigation into the nature of modality congruence in dual-task paradigms. Clearly, the current findings establish that attention must be paid to future methodologies utilizing a dual-task paradigm and the implications of modality congruence between the primary and secondary task. Furthermore, previous investigations that employ incongruent modalities between exposure and recognition must be revisited (Dewald et al, 2011, Sinnett et al, 2006).

Combined, the present findings and previous research offer insight into how irrelevant information is processed when it is presented simultaneously with an attended target. Under certain circumstances, as demonstrated here (see also Dewald et al., 2011 Dewald & Sinnett, submitted; Seitz & Watanabe, 2003; Tsushima et al., 2008), unattended stimuli can be perceived and affect behavior. Here we extend findings from previous research into the auditory sensory modality, and show a facilitation (i.e., attentional boost) for a highly exposed stimulus that was aligned with a target in the previous task when compared with items that were not aligned. More importantly however, we demonstrate that careful consideration must be given to the modality of presentation of dual-task paradigms in general.

References

Ahissar, M., & Hochstein, S. (1993). Attentional control of early perceptual learning. Proceedings of the National Academy of Science U.S.A, 90, 5718–5722.

Borst, A., & Egelhaaf, M. (1989). Principles of visual motion detection. *Trends in Neurosciences, 12*(8), 297-306.

Broadbent, D.E. (1958). *Perception and communication.* London: Pergamon Press.

Chandra., M., Robinson, C. W., & Sinnett, S. (2011). Coexistence of multiple modal dominances. In L. Carlson, C. Hölscher, & T. Shipley (Eds.), *Proceedings of the 33rd Annual Conference of the Cognitive Science Society*

Cherry, E.C. (1953). Some experiments on recogniton of speech, with one and two ears. *Journal of the Acoustical Society of America,* 25, 957-979.

Dewald, A.D., & Sinnett, S. (submitted). A window of perception when diverting attention? Enhancing recognition for explicitly presented, unattended, and irrelevant visual stimuli by target alignment. *Submitted to the Proceedings of the 34th Annual Conference of the Cognitive Psychology Society.*

Dewald, A.D., & Sinnett, S. (2011a). An inhibited recognition performance for explicitly presented target-aligned irrelevant stimuli in the auditory modality. *Proceedings of the 33rd Annual Conference of the Cognitive Psychology Society.*

Dewald, A.D. & Sinnett, S. (2011b). A multimodal investigation of recognition performance for target-aligned but irrelevant stimuli. *Proceedings of the 33rd Annual Conference of the Cognitve Science Society.*

Dewald, A.D., Sinnett, S., & Doumas, L.A.A. (2011). Conditions of directed attention inhibit recognition performance for explicitly presented target-aligned irrelevant stimuli. *Acta Psychologica, 138*, 60-67.

Duncan J, Martens S, Ward R (1997), Restricted attentional capacity within but not between sensory modalities. *Nature* 387(6635):808-10

Driver, J., & Spence, C. (2004). Cross-modal spatial attention: Evidence from human performance. In C. Spence & J. Driver (Eds.), Cross-modal space and cross-modal attention. Oxford, UK: Oxford University Press.

Lin, J.Y., Pype, A.D., Murray, & Boynton, G.M. (2010). Enhanced memory for scenes presented at relevant points in time. *PLoS Biol, 8*(3), E1000337.

Rees, G., Russell, C., Frith, C. D., & Driver, J. (1999). Inattentional blindness versus inattentional amnesia for fixated but ignored words. *Science, 286*, 2504-2507.

Roelfsema, P. R., van Ooyen, A., & Watanabe, T. (2009). Perceptual learning rules based on reinforces and attention. *Trends in Cognitive Sciences, 14*(2), 64-71.

Seitz, A. R., Kim, R., & Shams, L. (2006). Sound facilitates visual learning. *Current Biology, 16*, 1422- 1427.

Seitz, A. R. & Watanabe, T. (2003). Psychophysics: Is subliminal learning really passive? *Nature, 422*, 36.

Seitz, A. R. & Watanabe, T. (2005). A unified model for perceptual learning. *Trends in Cognitive Science, 9* (7), 329-334.

Sinnett, S., Costa, A., & Soto-Faraco, S. (2006). Manipulating inattentional blindness within and across sensory modalities. *Quarterly Journal of experimental Psychology, 59*(8), 1425-1442

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory, 6,* 1 74–215.

Swallow K. M., & Jiang, Y. V. (2010). The attentional boost effect: Transient increases in attention to one task enhance performance in a second task. *Cognition*, 115, 118-132.

Swallow K.M., & Jiang, Y. V. (2011). The role of timing in the attentional boost effect. *Attention, Perception, and Psychophysics,* 73, 389-404.

Toro, J.M., Sinnett, S., & Soto-Faraco, S. (2005). Speech segmentation by statistical learning depends on attention. *Cognition*, 97, 25-34

Triesman, A.M. (1964). Selective attention in man. *BritisishMedical Bulletin,* 20, 12-16.

Tshushima, Y., Sasaki, Y., & Watanabe, T. (2006). Greater disruption due to failure of inhibitory control on an ambiguous distractor. Science, 314, 1786-1788.

Tsushima, Y., Seitz, A. R., & Watanabe, T. (2008). Task-irrelevant learning occurs only when the irrelevant feature is weak. Current Biology,18(12), 516-517.

Watanabe, T., Náñez,Y., & Sasak, S. (2001). Perceptual learning without perception. *Nature, 413*, 844–848.

Wilson, M. D. (1988). The MRC psycholinguistic database: Machine readable dictionary, version 2. *Behavioural Research Methods, Instruments and Computers*, *20*, 6-11.