The influence of eye movements and illumination on auditory localization

GEORGE R. MASTROIANNI
University of New Hampshire, Durham, New Hampshire

A critical analysis of a recent paper by Shelton and Searle (1980) on the visual facilitation of auditory localization is presented. The author claims that Shelton and Searle fail to make the relevant distinction between Warren's (1970) frame of reference hypothesis and Jones's (1975) spatial memory hypothesis. Shelton and Searle's claim that they have demonstrated the existence of two distinct forms of visual facilitation is questioned. Data from an experiment in which auditory localization was tested under two levels of illumination (light and dark) and with two kinds of eye movement instructions (fixed and movement) are presented. The results show that facilitation occurs only when eye movements take place in a lighted environment. This is interpreted as supporting Warren's frame of reference hypothesis.

Over the past 10 years, several experiments have been conducted that have demonstrated the facilitating effect of visual stimulation on auditory localization (Jones, 1975; Jones & Kabanoff, 1975; Platt & Warren, 1972; Shelton & Searle, 1980; Warren, 1970). The most recent article in this series, that by Shelton and Searle (1980), discusses this literature in terms of the differing hypotheses advanced by the principal investigators in this field, Jones and Kabanoff (1975) and Warren (1970). Shelton and Searle claim that the differences in the hypotheses offered by these researchers are due to the differing nature of the processes tapped by each of their respective procedures. They present data from several experiments in an attempt to validate their view of the literature.

I will demonstrate in this paper that (1) their interpretation of the literature overlooks important contradictions in the empirical findings in this literature, (2) they misconstrue the nature of the hypotheses offered by Jones (1975) and Warren (1970), and (3) the experiments reported by Shelton and Searle (1980) do not, in fact, support their view of the literature. Finally, I will offer an alternative interpretation of the literature that more critically assesses the earlier findings and will present some data that were collected in an attempt to clarify some of the issues in this literature.

The literature on intersensory facilitation is riddled with contradictory findings and with failures to replicate earlier findings. A recent example is Platt and Warren's (1972) finding that having a subject make eye movements toward unseen auditory targets in the

dark resulted in no less variable or more accurate localizations than when subjects made no eye movements in the dark. Jones and Kabanoff (1975) reported that a condition that included eye movements in the dark did lead to facilitation relative to one in which vision was fixated in the dark. This particular contradiction is interesting because it serves to at least partially explain the theoretical differences between Jones (1975) and Warren (1970). Warren's data show that eye movements must occur in the light to produce facilitation, which is consistent with a frame of reference hypothesis. According to Warren's hypothesis, auditory localizations are more accurate in a lighted environment because eye position is known more accurately in the light, allowing a more precisely controlled pointing response. Thus, a textured, lighted environment provides a frame of reference for judgments of the position of auditory stimuli. Jones's results, showing that light is not a necessary condition, are more consistent with his memory stabilization hypothesis. According to this hypothesis, the "essential spatial framework would be given by motor activity." In particular, eye movements serve to update position memory; since eye movements are thought to be "noisy" in the dark, auditory localization is better in the light because eye drift—and, hence, auditory position memory—is stabilized. Shelton and Searle (1980) claim that the differences in the hypotheses advanced by Jones (1975) and by Warren (1970) are due to the different processes each is tapping. It is claimed here that a difference in assumptions based on a contradictory finding is at least partially responsible for these differences.

Shelton and Searle (1980) seem to think that the necessity for eye movements can be seen as a distinguishing feature of the two theoretical formulations. They state, "Jones and Kabanoff argued that the in-

Requests for reprints should be sent to George Mastroianni, P.O. Box 62, USAARL, Ft. Rucker, Alabama 36362. I would like to acknowledge the assistance of Kenneth Fuld and of Kathleen O'Donnell, whose discussions and comments were invaluable in clarifying some of the issues presented in the paper.

teraction was not explicable by a frame of reference hypothesis and that the data indicate the importance of eye movements to visual facilitation" (p. 590). The title of Platt and Warren's (1972) article, "Auditory Localization: The Importance of Eye Movements and a Textured Visual Environment," shows that Warren also believes eye movements to be necessary to the occurrence of visual facilitation. Indeed, according to Warren, eye movements in a textured visual environment calibrate or organize auditory space. The fundamental difference between Jones's (1975) and Warren's (1970) hypotheses, it seems to this author, is that Jones believes that sensory space is primarily organized by motor responses that serve to update position memory, rather than by information from any single modality such as vision. According to this conception, "visual facilitation" is really a misnomer; "motor facilitation" might be a more descriptive term. While Jones argues that eye movements per se are sufficient to produce facilitation, Warren holds that eye movements must occur in a textured visual environment to have this effect. Shelton and Searle seem to base their analysis on Warren's (1970) paper and, in fact, do not refer to Platt and Warren (1972). I would claim that the more recent paper contains the most complete statement of Warren's position.

The empirical data reported by Shelton and Searle (1980) are really only partially relevant to the facilitation literature. Both Jones (1975) and Warren (1970) used unseen auditory sources, while Shelton and Searle apparently made no attempt to occlude the speakers from the subject's view. Some of their speaker positions were unseen by virtue of their being out of range of peripheral vision. It is interesting to note that less facilitation was observed with unseen speakers than with speakers in visible locations. The visibility of the sound sources is an important issue because it transforms the nature of the subject's task. With unseen speakers, the subject's task is essentially the same with or without vision, but, with the visible speakers, the subject gains extra information in the light. That is, a set of discrete potential locations of sound origin is available to the subject. Small errors in auditory localization in the light can be overcome by assigning the sound to the nearest visible location, whereas the subject has no such cues to possible location in the dark. Under these circumstances, better localization would be expected as a matter of course with vision available and would not necessarily reflect a heightened sensitivity, as we believe visual facilitation causes.

The second part of Shelton and Searle's (1980) analysis was an "attempt to differentially control the two forms of visual facilitation." In an extremely complicated experiment, they observed that auditory localization was facilitated: (1) when a sound was

preceded by a pair of lights that were .3 deg nearer the midline than either of the two potential sound positions, but not when the sound was preceded by lights that were either cospatial with or .4 deg beyond the potential sound sources; and (2) when a sound was followed by lights that were cospatial with the potential sound positions, but not when the lights were either nearer to or farther from the midline than the sound positions. Shelton and Searle claim that these results support their position that two independent forms of visual facilitation exist—one dependent on frame of reference [(1) above], the other on spatial memory [(2) above]. According to Shelton and Searle, Warren (1970) found facilitation due to the frame of reference mechanism, while Jones's (1975) results were an example of facilitation due to the spatial memory mechanism.

If we are to accept Shelton and Searle's (1980) data as showing two separate mechanisms of facilitation, of which Jones's (1975) and Warren's (1970) results are separate instances, then the set of conditions under which Shelton and Searle found facilitation effects must not exclude the conditions under which Jones and Warren found their original results. To take the case of Warren, alleged to be tapping the frame of reference mechanism, the relevant Shelton and Searle result is the facilitation in the "preceding" condition with light pairs that were .3 deg nearer the midline. No facilitation was observed with light pairs that were cospatial with or .4 deg farther from the midline. Platt and Warren (1972) and Warren (1970) demonstrated facilitation with light that was spatially diffuse and temporally constant. That is, the illumination was general room illumination and was constantly on during the light conditions and constantly off during the dark conditions. If we are willing to make the questionable assumption that such constant illumination can function in the same way that the .3-deg lights do, but not as the .4-deg lights or the cospatial lights do, then we can accept Warren's results as an instance of Shelton and Searle's frame of reference mechanism. However, Jones (1975) also used continuous general room illumination and also found facilitation, due to the spatial memory mechanism, according to Shelton and Searle. This is analogous to the trailing condition of Shelton and Searle, when the lights followed the sounds. In this case, facilitation was observed only with the cospatial lights; none was observed with either the lights that were .3 deg nearer or those that were .4 deg farther from the midline. We are thus left in the position of having to assume that continuous general illumination functions like cospatial lights that follow an auditory stimulus, but also like lights .3 deg nearer the midline that precede an auditory stimulus. The conflicting nature of the assumptions that are required if we accept Shelton and Searle's re-

sults at face value call into question the claim that they involve the same mechanisms that were operative in Jones's and Warren's experiments.

I stated earlier that the differences between the theoretical formulations of Jones (1975) and those of Warren (1970) were at least partly due to contradictory empirical findings. The crucial difference in the empirical findings is that, in Jones's experiments, eye movements made in the dark facilitate localization performance relative to a condition in which the eyes are fixated in the dark. In the experiments of Warren and his colleagues, there is no significant difference between the two eye-movement conditions in the dark. If Jones had found the results that Warren did, then Jones would have had to abandon or modify his hypothesis about the mechanism of facilitation. Similarly, Warren's hypothesis is untenable if we believe, with Jones, that light is not necessary for facilitation to occur. The contradiction may be due to the fact that Jones and Warren used very different methods, dependent measures, and stimuli. Before such a conclusion can be drawn, however, it is necessary to demonstrate that both sets of findings can be reliably reproduced under similar experimental conditions by independent investigators. In my estimation, this has not occurred for either set of results (Mastroianni, 1981). Consequently, it would be both premature and uneconomical to accord the results of Jones and those of Warren the status of established findings on the basis of the relatively small number of data collected by each author. In particular, to conclude that both hypotheses must be correct and, therefore, that at least two different kinds of facilitation exist is inappropriate at the present stage of development of this literature, I believe.

In order to take a first step in providing the sort of replication that I feel is necessary, I conducted an experiment that was similar to Warren's (1970) experiments.

METHOD

Subjects

The subjects were 28 undergraduates at the University of New Hampshire who participated in the experiment to fulfill a course requirement.

Apparatus

Subjects were seated in a sound-treated room at a booth upon which a chinrest was mounted. The chinrest was placed at the center of a semicircle of 24-in. radius. Eight 1.5-in. Archer speakers were suspended on dowels along the perimeter of the circle at 10-deg intervals from 0 to ± 40 deg (0 being the midline) at approximately eye level. The speakers were hidden from view by a gray cloth screen that hung down along the perimeter of the circle. A paper screen ruled at 1-deg intervals hung down below the perimeter of the circle along which the speakers were located. A shelf extending out from the chinrest prevented the subject from seeing either this paper screen or his/her arm and hand. The subject held a 10-in. length of 1/8-in.-diam wooden dowel, which had a 28-V indicator light attached to its end. Subjects made localization responses by pointing the rod in the direction of the sound.

The experimenter, seated on the other side of the booth, could read the direction of the response by observing where the spot of light fell on the paper screen. Light conditions were manipulated by turning off the room lights and adding a blindfold to create the dark condition. The blindfold prevented the subject from being aware of the light produced by the small penlight that the experimenter used to read the response in the dark condition. In the light condition, room lights were left on and the subject wore no blindfold.

Auditory stimuli were produced by a white-noise generator. Their duration was controlled by an interval timer and maintained at 6 sec. Their intensity was 50 dB SPL; individual potentiometers for each speaker allowed matching to within ± 2 dB across the eight speakers.

Procedure

The independent variables were environmental illumination (light or dark) and eye-movement instructions (movement or fixation). Crossing these variables yielded four conditions: light movement, light fixation, dark movement, and dark fixation. In the movement conditions, subjects were instructed to look toward the sound when it came on and to keep looking at it until it went off, while keeping their heads straight ahead. In the fixated condition, subjects were instructed to maintain straight-ahead fixation throughout the procedure and were specifically instructed not to look in the direction of the sound when it came on.

The 28 subjects were each run in one light and one dark condition. Half of the subjects were run in light movement and half in light fixation. Of those run in light movement, half were run in dark movement and half in dark fixation. The same is true for those run in light fixation. Order of presentation of conditions was counterbalanced across subjects.

RESULTS

Two measures of responding were calculated for each condition. These were the mean of the errors of responses and the standard deviation (SD) of the error of the responses. Error was defined as the absolute value of the difference between the direction of the subject's response and the actual location of the speaker.

Two separate analyses of variance were performed. One, a 2 (illumination conditions) $\times$ 2 (eye-movement instructions) ANOVA was performed on the SD of error scores. There were no significant effects in this analysis. A second 2 $\times$ 2 ANOVA, this one on the mean error scores, showed a significant F value for the two-way interaction. Neither of the main effects was significant. A summary table for both analyses is presented in Table 1. An analysis of the simple main effects was also carried out. Three of the four simple effects were significant; the effect of illumination condition when subjects' eyes were fixated was not significant. It is interesting to note that the comparison of movement conditions at the dark level of the illumination variable was significant, but that the difference is in the direction opposite to what one might predict. That is, a higher average error is evident in the movement condition than in the fixated condition, when the environment is dark. On the basis of Jones's (1975) results, we would predict that error would be higher in the fixated condition. Given the relatively small sample size

Table 1
Analyses of Variance for Both Dependent Measures

Source	df	MS	F
Analysis of Standard Deviations of Responses			
Movement Instructions	1	13.85	2.75*
Light Condition	1	14.99	2.99*
Interaction	2	15.11	3.00*
Error	52	5.03	
Analysis of Mean Error of Responses			
Movement Instructions	1	.12	.04*
Light Condition	1	2.45	.77*
Interaction	2	11.29	3.53†
Error	52	3.20	

*Nonsignificant. † $p = .05$.

Table 2
Analysis of Simple Effects of Mean Error Scores

Source	df	MS	F
Movement Instructions in Light	1	12.96	4.05*
Movement Instructions in Dark	1	9.73	3.04*
Light Conditions for Fixated Vision	1	5.08	1.59†
Light Conditions for Eye Movements	1	19.96	6.24*

* $p = .05$. †Nonsignificant.

Table 3
Mean Error Scores

Movement Condition	Light Condition	
	Light	Dark
Fixated	4.47	4.12
Eye Movements	3.61	5.30

and high variability of these data, clarification of this anomaly must await further research. A summary table for the simple main effects is presented in Table 2. The table of means for the mean error scores can be found in Table 3.

DISCUSSION

The results reported here seem to support Warren's (1970) frame of reference hypothesis more than Jones's (1975) memory stabilization hypothesis. While Jones and Kabanoff (1975) do say that eye movements are likely to be more accurate in the light than in the dark and hence can stabilize auditory position memory better in the light than in the dark, their own results seem to imply that eye movements need not take place in an illuminated environment in

order to have a facilitating effect on localization. The results of both their Experiments 2 and 3 show just such an effect. This is consistent with their hypothesis that it is the motor responses that serve to update auditory position memory, and not any visual information gained during the eye movement. Warren's frame of reference hypothesis seems to this writer to require that eye movements take place in a lighted environment in order to produce facilitation. This difference can be seen as a way of experimentally distinguishing between the two hypotheses. On the other hand, the results reported here cannot be construed as a contradiction of Jones's results. The method I used was very different from Jones's, and this may have caused the differences observed.

The results I have reported here provide supporting evidence for the reliability of Warren's (1970) results. More such research, replicating both Jones's (1975) and Warren's studies should be undertaken before we consider either result to be valid. Only after several independent researchers have replicated the results of Jones and Warren should we seriously consider the theoretical issues here. If both findings are established as reliable, then parsimony would demand that we look for procedural differences that might allow us to reconcile the two sets of data and provide one general explanation. If the data still refuse to yield to attempts at reducing them to one or the other mechanism operating under varying sets of conditions, then, and only then, should we feel comfortable with a two-factor model, if the two hypotheses are intended to be general in scope and to include a variety of experimental situations.

REFERENCES

- JONES, B. Visual facilitation of auditory localization in school-children: A signal detection analysis. *Perception & Psychophysics*, 1975, 17, 217-220.
- JONES, B., & KABANOFF, B. Eye movements in auditory space perception. *Perception & Psychophysics*, 1975, 17, 241-245.
- MASTROIANNI, G. *The influence of visual texture on the visual facilitation of auditory vocalization*. Unpublished doctoral dissertation, University of New Hampshire, May 1981.
- PLATT, B., & WARREN, D. Auditory localization: The importance of eye movements and a textured visual environment. *Perception & Psychophysics*, 1972, 12, 245-248.
- SHELTON, B., & SEARLE, C. The influence of vision on the absolute identification of sound-source position. *Perception & Psychophysics*, 1980, 28, 589-596.
- WARREN, D. Intermodality interactions in spatial localization. *Cognitive Psychology*, 1970, 1, 114-133.

(Manuscript received May 21, 1981;
revision accepted for publication January 22, 1982.)