ELSEVIER

Acta Psychologica 108 (2001) 21–33

acta psychologica

www.elsevier.com/locate/actpsy

Directing spatial attention towards the illusory location of a ventriloquized sound

Jean Vroomen [a,*], Paul Bertelson [b], Beatrice de Gelder [b]

[a] *Department of Psychology, Tilburg University, Warandelaan 2, P.O. Box 90153, 5000 LE Tilburg, Netherlands*

[b] *Université Libre de Bruxelles, Belgium*

Received 12 July 2000; received in revised form 28 September 2000; accepted 30 September 2000

Abstract

In this study, we examined whether ventriloquism can rearrange external space on which spatial reflexive attention operates. The task was to judge the elevation (up vs down) of auditory targets delivered in the left or the right periphery, taking no account of side of presentation. Targets were preceded by either auditory, visual, or audiovisual cues to that side. Auditory, but not visual cues had an effect on the speed of auditory target discrimination. On the other hand, a ventriloquized cue, consisting of a tone in central location synchronized with a light flash in the periphery, facilitated responses to targets appearing on the same side as the flash. That effect presumably resulted from the attraction of the apparent location of the tone towards the flash, a well-known manifestation of ventriloquism. Ventriloquism thus can reorganize space in which reflexive attention operates. © 2001 Elsevier Science B.V. All rights reserved.

PsycINFO classification: 2320

Keywords: Ventriloquism; Attention; Cross-modal perception

Perceptual processing has generally been studied in one specific sense modality at a time. Yet, most events in real life produce stimulation that impinges simultaneously

* Corresponding author. Tel.: +31-13-4662394.
E-mail address: j.vroomen@kub.nl (J. Vroomen).

PII: S 0 0 0 1 - 6 9 1 8 (0 0) 0 0 0 6 8 - 8

on several modalities. Evidence that the system somehow combines corresponding inputs from separate modalities has come mainly from studies with conflict situations in which intermodal discordances are created on some dimension of the inputs, like location or identity, while other dimensions are kept coherent. A well-known example is the ventriloquist situation where synchronous auditory and visual events are presented in somewhat separate locations. One of the main manifestations of ventriloquism is *immediate cross-modal bias*. When subjects are asked to indicate, by pointing or by some kind of verbal response, the location of the auditory input and to ignore the spatially discordant visual input, the reported location of the target sound is displaced in the direction of the visual input (Bertelson, 1994, 1999; Bertelson & Radeau, 1981; Radeau, 1994; Radeau & Bertelson, 1978; Welch, 1978).

A central question about spatial cross-modal bias concerns the locus of the underlying processes in the cognitive architecture. The usual demonstration of the phenomenon in the selective localization task consists in a partial failure to follow selective modality instructions. On the one hand, this may reflect a mandatory operation of a process that integrates spatial information across input modalities. Yet, it may also be the case that voluntary post-perceptual decisions contribute. In the typical selective localization task, subjects are left free to speculate why to respond only to the auditory modality and not the visual, and they may adopt response strategies different from instructions (Bertelson, 1999; Vroomen, 1999). For example, subjects may, on some trials, point by mistake towards the visual input. They may also deliberately adjust their response strategy towards a compromise solution of the conflicting auditory and visual inputs. In a typical ventriloquist task, it therefore seems likely that many levels of processing will contribute to the actual performance.

Previously, we have argued that at least part of the ventriloquist phenomenon reflects an automatic perceptual process (Bertelson, 1999; Bertelson & Aschersleben, 1998; Vroomen, 1999). For example, we have shown that ventriloquism is not mediated by where *endogenous* visual attention is focussed (Bertelson, Vroomen, de Gelder & Driver, 2000). When subjects had to localize auditory target sounds and ignore bright visual distracters flashed synchronously to the left or the right of the sounds, it did not matter whether they were focussing on the peripheral distracter rather than a central location. Equal amounts of ventriloquism were obtained in the two cases. In another set of experiments, we have shown that ventriloquism is not influenced by whether or not the visual attracter receives *exogenous* visual attention (Vroomen, Bertelson & de Gelder, in press). A display was used that consisted of a row of four bright squares with one square, in either left- or right-most position, *smaller* than the others, serving as singleton. It appeared that the small singleton effectively captured exogenous attention, but that the apparent location of a target sound was instead shifted towards the bigger squares. Taken together, these results indicate that ventriloquism is an automatic process that is largely independent from endogenous and exogenous spatial attention.

In the present study, we adopted a new methodological approach to the same issue by measuring the effect of a *ventriloquized* sound upon processing of a subsequent stimulus. The main advantage of such an indirect approach is that voluntary response strategies are unlikely to play a role. If the ventriloquized location of an auditory cue is nevertheless found effective, then the result can more safely be attributed to a genuine perceptual process. We used the cross-modal cueing task introduced by Spence and Driver (1997) in which auditory targets are presented in either the left or the right periphery, and in one of two elevations, up or down. The task is to judge the elevation while ignoring azimuthal location, but the latter is cued by an uninformative spatially congruent or incongruent stimulus in one of several modalities. Spence and Driver (1997) obtained cueing effects (i.e., faster responses with congruent than with incongruent cues) across all modalities, except that *visual cues do not affect responses to auditory targets* (see Driver & Spence, 1998 for an overview). Our question was what happens with a *ventriloquized* sound, i.e., one whose physical source is in the centre, but whose apparent source is shifted laterally by a light presented simultaneously on the left or right. If such a ventriloquized sound attracts exogenous attention to its illusory location, it would suggest that cross-modal integration (i.e., ventriloquism) precedes the orienting of spatial attention, or alternatively, that it co-occurs at approximately the same time. In both cases, though, ventriloquism would rearrange space on which attention operates.

Unknown to us, Spence and Driver (2000) have independently conducted an experiment along these general lines. As a cue, they used a sound that could be ventriloquized in the vertical plane by a simultaneous light. The elevation of the sound was either easy or hard to localize unimodally (i.e., white noise vs a pure 2000 Hz tone, respectively). They obtained cueing effects with hard-to-localize sounds, but not with easy-to-localize ones, a difference they attributed to a higher susceptibility of poorly localized sounds to cross-modal interference. Cueing effects of the hard-to-localize sounds (an 11 ms effect) were significant at the shortest stimulus onset asynchrony (SOA) (100 ms), but not at longer SOAs (200 and 700 ms).

The present study is similar to Spence and Driver (2000), although we used ventriloquism in the horizontal plane rather than in the vertical one, because it is the dimension in which ventriloquism has been studied most often. SOAs were also somewhat different (100, 300 and 500 ms). We first checked (in Experiment 1) that an auditory cue, but not a visual one, affects auditory target discrimination. We also used an *audio-visual* cue consisting of a light flashed *at the same location* as the sound. It seemed important to determine whether such a bimodal cue had any effect that was not predicted by its unimodal components. In neuro-anatomical studies, it has indeed been found that some multi-modal cells of cat's superior colliculi show a multiplicatively enhanced response when a sound and light are presented together in their receptive field (Stein & Meredith, 1993). Since it has been argued that the findings from the cross-modal cueing paradigm are related to the behavior of these cells (Driver & Spence, 1998), one might expect that sounds and lights presented simultaneously in the same location would exert similarly enhanced cueing effects.

1. Experiment 1

1.1. Method

1.1.1. Subjects

Nineteen subjects were recruited by advertisement to take part in this experiment. All were first year students from Tilburg University and they received course credits for their participation. Data of three subjects were discarded from the analysis because their performance in auditory elevation discrimination was below 80% correct. The remaining subjects, 13 females and 3 males, all had normal hearing and vision.

1.1.2. Apparatus and materials

All experiments were conducted in a dimly lit soundproof booth. Subjects were seated at a black table facing straight ahead with head movements precluded by a chinrest with cheeck pads. A small green light emitting diode (LED) was placed 53 cm in front of the subject at eye level and served as fixation point. Eye movements were monitored by the experimenter via a hidden camera.

The auditory cues were presented via either of two loudspeakers (Philips box 410 Car loudspeakers, 30 W with a cone diameter of 9 cm) mounted on an adjustable rack at ear level, 44° from fixation (see Fig. 1 for a schematic view of the setup). The

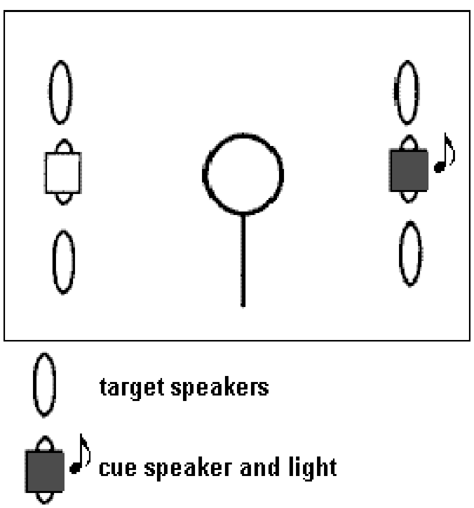

Fig. 1. A schematic outline of the setup of Experiment 1 as seen from behind the subject's head. Audiovisual cues consisted of a tone and light presented from the same location. Not shown is that in Experiments 2 and 3, the tone for a ventriloquist cue was delivered from a centrally located hidden loudspeaker.

visual cues were presented by either of two sets of three red LEDs enclosed in a circular cover of reflecting metal, and mounted in front of the loudspeakers that delivered the auditory cues. The diameter of the cover was 5.2 cm, and the eccentricity was also 44° from fixation. The luminance of the LEDs was 40 cd/m^2. A visual cue was presented on one side by illuminating all three LEDs in one of the covers. More than one LED was used to ensure that the cues were sufficiently salient to produce covert orienting.

The auditory targets were created by a white-noise generator, and were presented from any one of four additional loudspeakers of the same type as those used for the cues. Two target loudspeakers were placed on either side of the subject, each arranged 23.5 cm above or below a cue loudspeaker. The auditory cue consisted of a 2000 Hz tone presented for 100 ms (with a 5 ms linear fade-in and fade-out) at 63 dB(A) when measured at the ear level. The sounds were designed to be unlocalizable in elevation, yet localizable in azimuth (see Spence & Driver, 1997). The target sound consisted of five 20 ms bursts of white noise at 69 dB(A), each separated by 10 ms silent gaps. These target sounds were designed to be localizable in both azimuth and elevation.

The up–down decision for the targets required a discrimination of 28° vertically. Subjects responded by pressing one of the two keys attached to the table top in front of them, one immediately behind the other. Response times (RTs) were measured in milliseconds from target onset by using a microcomputer with a special card to interface the loudspeakers, the LEDs, and the key responses.

1.1.3. Design

There were eight blocks of 144 experimental trials each. Each block was preceded by four warming-up trials. There were three within-subjects factors: modality of the cue (auditory, visual, or audio-visual), the SOA between the cue and subsequent target (100, 300, and 500 ms), and whether the cue was on the same side as the target or on the opposite side (valid vs invalid trials, respectively). When crossed with the four possible target locations (left or right, up or down), these factors yielded 36 equiprobable conditions, each occurring four times pseudo-randomly within each block.

1.1.4. Procedure

The fixation light was illuminated at the beginning of each trial and remained on until a response was made. Participants were instructed to maintain fixation on this central green LED whenever it was illuminated. After a random delay of 400–650 ms, a cue (unpredictably in the auditory, visual, or audio-visual modality) was presented (cue side was unpredictably on the left or right) for 100 ms. After a further delay (unpredictably 0, 200, or 400 ms) the auditory target was presented (unpredictably on the same side of the cue or in the opposite direction).

Subjects pressed with their right-middle finger the key farthest from them for a target from the upper position (regardless of the side), and pressed with their right-index finger the nearest key for a target from the lower positions (regardless of side). To give subjects feedback concerning their performance, the fixation light was turned off immediately following a correct response, but flickered for 150 ms following an erroneous response. Once extinguished, the fixation light remained off for 1 s before coming on again to signal the start of the next trial. Subjects were told that the cues were completely uninformative since targets were equally likely to come from the side opposite to the cue as from the same side.

1.2. Results

The first block was treated as practice, and was therefore discarded. Less than 1% of the data was removed because of eye movements. Incorrect responses and warming-up trials were also discarded from the analysis. The overall error rate in this and all the other experiments was quite low (it never exceeded 6.1%) if compared with Spence and Driver (1997, Experiment 1) in which there was an overall error rate of 16.8% for auditory targets. In subsequent experiments, there was also no speed-accuracy trade-off in the critical ventriloquist conditions, so error rates were not analyzed any further.

RT data were pooled across all locations of the target (up–down and left–right). Median RTs of each subject were calculated for each combination of cue modality (visual, auditory, audio-visual), cue validity (valid, invalid), and SOA. Means of these scores across subjects are shown in Table 1, together with the error rates. A three-way within-subject ANOVA yielded main effects of validity, $F(1,15) = 52.74$, $P < 0.001$ (faster responses on valid than on invalid trials), and of SOA, $F(2,30) = 76.19$, $P < 0.001$ (slower responses at the shortest SOA). There were, furthermore, significant interactions between modality and SOA, $F(4,60) = 8.24$, $P < 0.001$, modality and validity, $F(2,30) = 39.50$, $P < 0.001$, SOA and validity, $F(2,30) = 16.04$, $P < 0.001$, as well as a significant second-order interaction between the three factors, $F(4,60) = 3.48$, $P < 0.02$.

Fig. 2 presents average cueing effects measured by the difference between median RTs on invalid trials and on corresponding valid trials. The effects of auditory and audio-visual cues appeared nearly identical, and were very different from those of

Table 1
Mean RTs (in ms) and error rates (as percentages, in parentheses) for auditory, visual, and audio-visual cues in Experiment 1

SOA (ms)		Auditory cue		Visual cue		Audio-visual cue	
		RT	Error	RT	Error	RT	Error
100	Valid	483	(5.0)	506	(4.5)	497	(6.9)
	Invalid	553	(6.0)	496	(2.9)	561	(7.8)
300	Valid	442	(3.6)	464	(3.7)	446	(4.6)
	Invalid	475	(3.9)	456	(3.0)	472	(2.8)
500	Valid	446	(3.3)	468	(3.0)	446	(3.3)
	Invalid	470	(2.8)	478	(3.0)	472	(2.1)

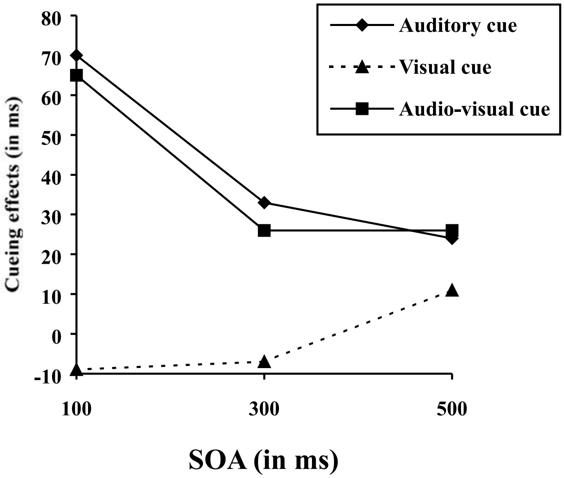

Fig. 2. Average cueing effects (in ms) of auditory, visual, and audio-visual cues as a function of SOA in Experiment 1.

visual cues. Separate ANOVAs with SOA as within-subjects factor were carried out on the data for the different cue modalities. *Visual cues* had no overall cueing effect $F < 1$, but there was a small significant effect of SOA, $F(2,30) = 3.88$, $P < 0.05$ (cueing effect increased at longer SOAs). In contrast, both *auditory* and *audio-visual cues* had a substantial overall effect (auditory cues, $F(1,15) = 49.29$, $P < 0.001$; audio-visual cues, $F(1,15) = 92.27$, $P < 0.001$), and the effect of SOA (auditory cues, $F(2,30) = 13.29$, $P < 0.001$; audio-visual cues, $F(2,30) = 20.50$, $P < 0.001$ indicating that the cueing effects *decreased* at longer SOAs. Pairwise *t*-test for the cueing effects at each SOA showed that the effects of audio-visual cues were never different from those of the auditory cues (all $P's > 0.20$), while the two latter effects were always significantly larger than those of visual cues (all $P's$ at least < 0.05).

1.3. Discussion

Experiment 1 confirms Spence and Driver (1997) findings that visual cues do not attract auditory attention. The effect of the audio-visual cue was not bigger than that of the auditory cue, showing that combining a sound and a light in the same location does not add to the cueing power of the sole sound. This result is important regarding the possible effects of ventriloquized cues to be examined in the following experiments. It means that an eventual effect cannot be interpreted as caused directly by the effect of combining two modalities.

In Experiment 2, subjects again made speeded responses to the elevation of auditory targets. In the critical condition, a target was preceded by a *ventriloquist cue* consisting of a central tone and a synchronous light flash in the same peripheral location as used before for the visual cues. The apparent location of the tone must normally be attracted towards the light, making it possible to determine if an illusory displacement can capture spatial attention.

2. Experiment 2

2.1. Method

Nineteen new subjects took part in this experiment. One subject was unable to make the auditory up–down discrimination above 80% and his data were therefore removed from the analysis. Of the remaining subjects nine were females and nine males.

The apparatus, materials, and design were exactly the same as in Experiment 1, except that cues from the auditory condition were replaced by ventriloquist cues. A ventriloquist cue consisted of a lateral visual cue synchronized with a cue tone emitted from a centrally located loudspeaker (of the same type as the other speakers) hidden behind a black curtain at a distance of 90 cm.

2.2. Results and discussion

Data were treated as in Experiment 1. The means of subjects' median RTs are shown in Table 2, separately for each modality of the cue and each SOA. A three-way within-subjects ANOVA was conducted on the median RTs with the factors cue modality (ventroloquist, visual, audio-visual), validity, and SOA. There was a main effect of validity, $F(1,17) = 26.31,\ P < 0.001$, (faster responses on valid than on invalid trials), and a significant effect of SOA, $F(2,34) = 115.20,\ P < 0.001$ (slower responses at the shortest SOA). There were, furthermore, significant interactions between cue modality and SOA, $F(4,68) = 19.48,\ P < 0.001$, cue modality and validity, $F(2,34) = 31.25,\ P < 0.001$, SOA and validity, $F(2,34) = 3.31,\ P < 0.05$,

Table 2
Mean RTs (in ms) and error rates (as percentages, in parentheses) for ventriloquist, visual, and audio-visual cues in Experiment 2

SOA (ms)		Ventriloquist cue		Visual cue		Audio-visual cue	
		RT	Error	RT	Error	RT	Error
100	Valid	527	(5.5)	516	(5.1)	519	(5.2)
	Invalid	520	(4.5)	517	(4.3)	595	(4.8)
300	Valid	469	(5.4)	498	(3.8)	466	(4.2)
	Invalid	478	(6.6)	492	(4.6)	507	(6.0)
500	Valid	468	(4.6)	510	(4.0)	469	(5.4)
	Invalid	482	(4.8)	506	(3.7)	486	(5.1)

and a significant second-order interaction between cue modality, validity, and SOA, $F(4,68) = 9.56,\ P < 0.001$.

Fig. 3 shows average cueing effects (again RT differences between invalid and valid trials). Separate ANOVAs of these data for each modality of the cue with SOA as within-subjects factor showed that *visual cues* had neither an overall effect nor an effect of SOA (both $F = s < 1$), while *audio-visual* cues had a substantial overall effect, $F(1,17) = 45.60,\ P < 0.001$ that decreased with longer SOAs, $F(2,34) = 12.05,\ P < 0.001$. The overall cueing effect of the *ventriloquist cue* was not significantly different from 0, $F(1,17) = 2.03,\ P = 0.17$, but there was an effect of SOA showing that the cueing effect *increased* with longer SOAs, $F(2,34) = 3.34,\ P < 0.05$. Separate t-tests (one-tailed) for the cueing effects at each SOA showed that at 100 ms SOA, audio-visual cues had a bigger effect than both visual cues, $t(17) = 5.85,\ P < 0.001$, and ventriloquist cues, $t(17) = 5.51,\ P < 0.001$. The two latter effects were not significantly different from each other ($P > 0.20$). At 300 ms SOA, audio-visual cues had bigger effects than both visual cues, $t(17) = 5.49,\ P < 0.001$, and ventriloquist cues, $t(17) = 4.45,\ P < 0.001$, and, more importantly, ventriloquist cues had bigger effects than visual cues, $t(17) = 2.05,\ P < 0.028$. At 500 ms SOA, audio-visual and ventriloquist cues both had bigger effects than visual cues, audio-visual: $t(17) = 1.95,\ P < 0.034$; ventriloquist: $t(17) = 2.38,\ P < 0.015$, and there was now no difference anymore between the audio-visual and ventriloquist cues ($P > 0.20$).

To sum up, while confirming the effectiveness of an audio-visual cue and the lack of effectiveness of a visual cue, Experiment 2 has supported the prediction that spatial attention can be drawn to the to the illusory location of a ventriloquized sound. The effect, however, was only significant at the longer SOAs of 300 and 500 ms. In Experiment 3, we tried to extend the exploration with a finer-grained range of SOAs.

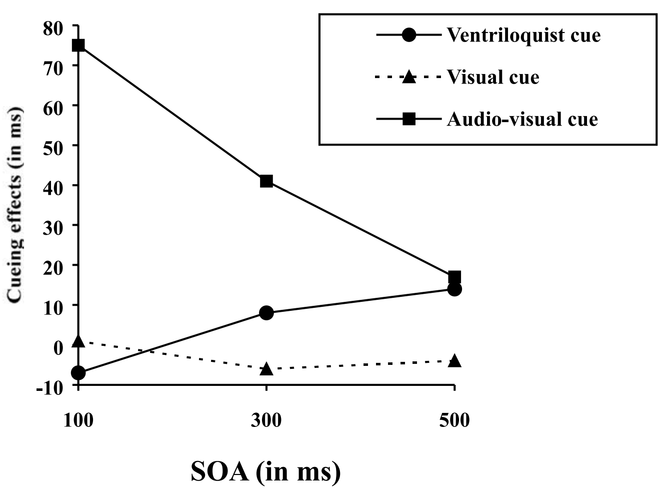

Fig. 3. Average cueing effects (in ms) of ventriloquist, visual, and audio-visual cues as a function of SOA in Experiment 2.

3. Experiment 3

3.1. Method

Twenty-six new subjects took part in this experiment. The apparatus, materials, and design were all as in Experiment 2. The same number of trials were run, but now evenly distributed across six SOAs (100, 200, 300, 400, 500, and 600 ms).

3.2. Results and discussion

The intersubject means of median RTs are shown in Table 3, separately for each modality of the cue and each SOA. In the overall ANOVA on the median RTs, there was a significant effect of SOA, $F(5,125) = 35.28,\ P < 0.001$ because responses were again slower at the shortest SOA. There were significant interactions between modality of the cue and SOA, $F(10,250) = 10.89,\ P < 0.001$, modality of the cue and validity, $F(2,50) = 60.43,\ P < 0.001$, SOA and validity, $F(5,125) = 6.64$, $P < 0.001$, as well as a significant second-order interaction between the three factors, $F(10,250) = 29.43,\ P < 0.001$.

Fig. 4 presents the average cueing effects (RT invalid trial–RT valid trials). Separate ANOVAs per modality with SOA as within-subjects factor showed that again *visual cues* had no overall cueing effect and no effect of SOA (both F's<1). *Audio-visual cues* had a substantial overall cueing effect, $F(1,25) = 66.09,\ P < 0.001$, which decreased with longer SOAs, $F(5,125) = 3.85,\ P < 0.003$. The overall cueing effect

Table 3
Mean RTs (in ms) and error rates (as percentages, in parenthesis) for ventriloquist, visual, and audio-visual cues in Experiment 3

SOA (ms)		Ventriloquist cue		Visual cue		Audio-visual cue	
		RT	Error	RT	Error	RT	Error
100	Valid	563	(9.3)	535	(7.6)	549	(9.4)
	Invalid	536	(6.1)	530	(3.9)	596	(8.0)
200	Valid	509	(7.5)	520	(6.1)	497	(9.2)
	Invalid	502	(4.9)	512	(4.8)	517	(6.0)
300	Valid	487	(6.7)	504	(6.1)	477	(6.3)
	Invalid	496	(6.4)	511	(4.8)	503	(5.2)
400	Valid	481	(6.7)	500	(4.6)	485	(7.0)
	Invalid	491	(5.6)	500	(5.3)	496	(5.6)
500	Valid	474	(4.9)	503	(5.4)	480	(7.1)
	Invalid	496	(7.0)	504	(6.7)	501	(3.5)
600	Valid	491	(5.7)	507	(5.7)	490	(5.6)
	Invalid	492	(5.2)	506	(4.2)	505	(5.2)

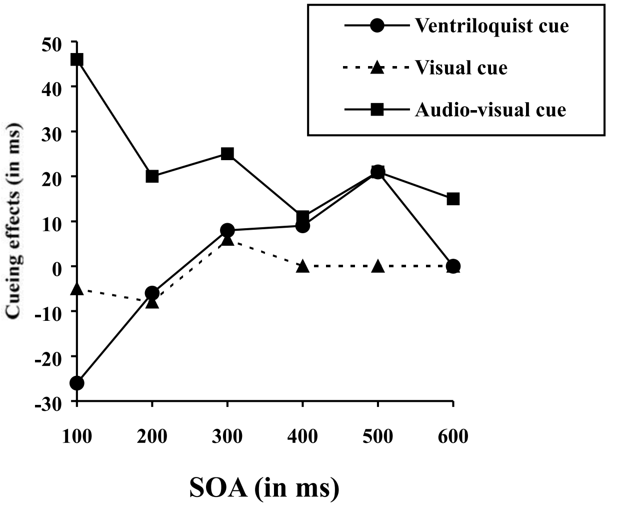

Fig. 4. Average cueing effects (in ms) of ventriloquist, visual, and audio-visual cues as a function of SOA in Experiment 3.

of the *ventriloquist cue* was not significantly different from 0, $F(1,25) < 1$, NS, but there was a significant effect of SOA, $F(5,125) = 5.89$, $P < 0.001$.

By t-tests (one-tailed), the effect of ventriloquist cues was significantly bigger than that of visual cues only at the 500 ms SOA, $t(25) = 2.71$, $P < 0.006$, while all other pairwise differences between these modalities were non-significant (all P's>0.05). Audio-visual cues had bigger cueing effects than ventriloquist cues at short SOAs (100 ms SOA: $t(25) = 6.53$, $P < 0.001$; 200 ms SOA: $t(25) = 3.10$, $P < 0.005$), but there was no difference anymore at SOAs longer than 300 ms (all P's>0.05). Finally, audio-visual cues had bigger cueing effects than visual cues at all SOAs (all P's>0.05).

Experiment 3 thus replicated the previous experiment in showing that ventriloquist cues have a cueing effect at around 500 ms SOA, and possibly somewhat earlier as well.

3.3. General discussion

The present study showed that a ventriloquized sound, i.e., a sound whose illusory location is shifted towards a synchronously presented light, could attract spatial attention to its illusory location. The result cannot be attributed to a direct cueing effect of the visual component of the ventriloquist cue, because (1) visual cues had no effect on auditory target discrimination, and (2) the effect of a single-location audio-visual cue (Experiment 1) was not superior to that of a unimodal auditory cue. The most likely explanation is therefore that the cueing effect of the ventriloquist cue emerges because the light component attracts the apparent location of the sound component, and it is the illusory location of the sound component that reflexively attracts attention in its direction.

One concern about the present data is that the temporal deployment of the ventriloquist cueing effect was different from that of an audio-visual cue. The ventriloquist cueing effect had a peak at around 500 ms, whereas audio-visual cues had their biggest effects at short SOAs. One way to account for this difference is that the apparent location of the ventriloquized cue is certainly less eccentric than that of the audio-visual cue, because there is abundant proof that the visual bias of auditory location is smaller than the distance between auditory target and visual distracter (Bertelson & Radeau, 1981; Radeau & Bertelson, 1978). The perceived location of a (valid) audio-visual cue is thus closer to the target than that of a ventriloquist cue, which explains why audio-visual cues are more effective than ventriloquist cues.

At the moment, it is not clear whether this difference in the perceived location of the audio-visual and ventriloquist cue also explains their different time courses. In contrast with our study, Spence and Driver (2000) obtained an effect on a ventriloquized sound at around 100 ms SOA, but not at 200, nor at 700 ms SOA (they did not measure the cueing effect at 500 ms). In their study, though, it seems that the illusory location of the hard-to-localize ventriloquized sound was close to the target itself (i.e., a big ventriloquist effect) because auditory information about vertical location was ambiguous. The ventriloquist effect in our experiments is probably smaller, because horizontal localization of a sound is, in general, more precise than vertical localization, and so the effect of the visual stimulus is expected to be of relatively less importance in the latter case. Whether this difference in the distance of the perceived location of targets and ventriloquized sound is responsible for the different time courses of the ventriloquist cueing effects in the two studies remains to be investigated.

Our finding that single-location audio-visual cues were not more effective than unimodal auditory cues in the same location is important not only because of its role in analyzing the mechanism of the ventriloquist cueing effect. Similar results were described in Spence & Driver (1999) who found that audio-visual cues were not more effective and, under some conditions, bimodal cues were even less effective than unimodal auditory cues. As explained in the introduction, one might have expected a stronger effect of audio-visual cues on the basis of the data concerning the response of multimodal cells in the cat's superior colliculus to similar multimodal compounds Stein & Meredith, 1993. Results of cross-modal cueing experiments have indeed been related to these neuro-anatomical facts (Driver & Spence, 1998). It seems, though, that this relation is not as straightforward as has been suggested, and that more research is needed, in particular with other multi-modal cues.

To conclude, this study showed that the auditory scene on which exogenous attention operates can be reorganized through ventriloquism. This suggests that cross-modal integration can arise preattentively, or in co-occurrence with the orientation of spatial attention. This is consistent with recent data showing that a sound can affect early stages of visual processing (Vroomen & de Gelder, 2000) and with the finding of Driver (1996) that it is easier to shadow a message when it is ventriloquized away from a distracter sound. It is also consistent with the recently demonstrated facts that ventriloquism does not depend on whether a visual attractor receives either endogenous (Bertelson et al., 2000) or exogenous attention (Vroomen et al., in press).

Acknowledgements

This work was supported by the Ministry of Scientific Research of the Belgian French-speaking Community (Concerted Research Action 96/01-2037) and by the Belgian National Fund for Collective Fundamental Research (Contract 2.45.39.95). We like to thank Esther IJpelaar for help in testing subjects.

References

Bertelson, P. (1994). The cognitive architecture behind auditory-visual interaction in scene analysis and speech identification. *Current Psychology of Cognition*, *13*, 69–75.

Bertelson, P. (1999). Ventriloquism: A case of cross-modal perceptual grouping. In G. Aschersleben, T. Bachmann, & J. Müsseler (Eds.), *Cognitive contributions to the perception of spatial and temporal events* (pp. 347–362). Amsterdam: Elsevier.

Bertelson, P., & Arhenleben, G. (1998). Automatic visual bias of perceived auditory location. *Psychonomic Bulletin & Review*, *5*, 482–489.

Bertelson, P., & Radeau, M. (1981). Cross-modal bias and perceptual fusion with auditory-visual spatial discordance. *Perception & Psychophysics*, *29*, 578–587.

Bertelson, P., Vroomen, J., de Gelder, B., & Driver, J. (2000). The ventriloquist effect does not depend on the direction of deliberate visual attention. *Perception & Psychophysics*, *62*, 321–332.

Driver, J. (1996). Enhancement of selective listening by illusory mislocation of speech sounds due to lip-reading. *Nature*, *381*, 66–68.

Driver, J., & Spence, C. J. (1998). Attention and the cross-modal construction of space. *Trends in Cognitive Sciences*, *2*, 254–262.

Radeau, M. (1994). Auditory-visual spatial interaction and modularity. *Current Psychology of Cognition*, *13*, 3–51.

Radeau, M., & Bertelson, P. (1978). Cognitive factors and adaptation to auditory-visual discordance. *Perception & Psychophysics*, *22*, 137–146.

Spence, C., & Driver, J. (1997). Audiovisual links in exogenous covert spatial attention. *Perception & Psychophysics*, *59*, 1–22.

Spence, C. & Driver, J. (1999). A new approach to the design of multimodal warning signals. In D. Harris (Ed.), *Engineering psychology and cognitive ergonomics, Vol. IV* (pp. 455–461). Aldershot: Ashgate Publishing.

Spence, C., & Driver, J. (2000). Attracting attention to the illusory location of a sound: Reflexive crossmodal orienting and ventriloquism. *Neuroreport*, *11*, 2057–2061.

Stein, B. E., & Meredith, M. A. (1993). *The merging of the senses*. Cambridge, MA: MIT Press.

Vroomen, J. (1999). Ventriloquism and the nature of the unity assumption. In G. Aschersleben, T. Bachmann, & J. Müsseler (Eds.), *Cognitive contributions to the perception of spatial and temporal events* (pp. 389–393). Amsterdam: Elsevier.

Vroomen, J., Bertelson, P., & de Gelder, B. (in press). The ventriloquist effect does not depend on the direction of automatic visual attention, *Perception & Psychophysics*.

Vroomen, J., & de Gelder, B. (2000). Sound enhances visual perception: Cross-modal effects of auditory organization on vision. *Journal of Experimental Psychology: Human Perception and Performance*, *26*, 1583–1590.

Welch, R. B. (1978). *Perceptual modification adaptation to altered sensory environments*. New York: Academic Press.